A CONTRAST BETWEEN VP-ELLIPSIS AND GAPPING IN ENGLISH:

L1 ACQUISITION, L2 ACQUISITION, AND L2 PROCESSING

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

SECOND LANGUAGE STUDIES

MAY 2020

By

Haerim Hwang

Dissertation Committee:
Bonnie D. Schwartz, Chairperson
Robert Bley-Vroman
Kamil Deen
William O'Grady
Amy Schafer

Keywords: VP-Ellipsis, Gapping, learnability, first language acquisition, early and late second language acquisition, adult second language processing

ACKNOWLEDGMENTS

I would like to thank everyone who supported me and participated in my efforts to write this dissertation. First and foremost, I sincerely appreciate my advisor, Bonnie D. Schwartz, who has guided me through this long and arduous journey with her vast knowledge, keen insight, and warm-hearted support. Over the past five years, she has become the person who I share the most treasured intellectual and emotional rapport that I have with. She is an awesome advisor, proactively listening to my thoughts, respecting my ideas, and challenging me with critical comments and questions. In particular, I am very grateful for her helpful feedback on my dissertation, which led me to think more critically and logically about the assumptions, interpretations, and conclusions that I had made. Having meetings with her and being in her classes were a genuine joy to me. They brought me new ideas and insights into the structure of language and second language acquisition plus tons of laughs as a bonus, thanks to her sense of humor. Thank you very much and I love you, Bonnie!

I would like to extend my gratitude to Robert Bley-Vroman, who kindly agreed to serve on my committee even after his retirement. I think I was very lucky to be a teaching assistant of SLS 441 in Fall 2017, when I was able to meet him from time to time in order to grade exams. The deep thinking and logical comments he shared throughout our discussions while grading made me feel brave enough to ask him to be my committee member—and he said yes! This dissertation is far more theoretically sound than it might have been thanks to his valuable input.

I was also fortunate to have Kamil Deen as one of my committee members. I truly admire his deep and broad knowledge, inspirational wisdom, and endless support. A part of this dissertation started from his language acquisition course. He generously donated his time to talk with me about my ideas at every stage of this dissertation project. His questions and comments helped me to keep focusing on the fundamental points of my dissertation and to take steps forward in my research. Thank you, Kamil, from the bottom of my heart for all your help.

I would like to sincerely thank William O'Grady for his invaluable input to this dissertation, which was always accompanied by kind words and encouragement. I always enjoyed thinking about the interesting language examples he shared with me and trying to address his comments and questions. He showed me how fun it can be to do what linguists do.

He made me dream of being a true scholar who takes on academic challenges with an open mindset and professional attitude, just like him.

Amy Schafer has had a profound influence on my research and thinking during my PhD studies, which I can never thank her for enough. She taught me how to conduct thorough and precise research both inside and outside of class. I really appreciate her for going over my audio stimuli for this dissertation, discussing their prosody with me, and giving me helpful questions and comments. Her detailed and insightful feedback and warm-hearted words always encouraged me to think further and try my best to find ways to improve this dissertation.

I am also grateful to other faculty members at the University of Hawai‘i. I deeply thank Theres Grüter for teaching me countless crucial things, including the importance of stepping back to see the bigger picture when necessary. Thanks to Shin Fukuda, I was able to learn how to make a stronger syntactic argument with independent evidence. I also appreciate Li ‘Julie’ Jiang and Shigeo Tonoike for facilitating my academic growth in syntax, Lyle Campbell and Patricia Donegan for training me in phonology, Kristopher Kyle and Rory Turnbull for introducing me to the world of programming, James Dean Brown for teaching me about curriculum development, and Gabriele Kasper for introducing me to various aspects of conversation analysis.

I would like to express my heartfelt thanks to Kenton Harsch and the SLS BA Committee for giving me an opportunity to teach undergraduate courses at the University of Hawai‘i. I am indebted to Kenton for his advice and guidance, which helped me to be a professional teacher and a warmer person. I want to add that I will very much miss him as well as the nice SLS undergraduate students that I met through teaching.

I greatly appreciate my professors in Korea. My special thanks go to Hyun-Kwon Yang for encouraging me to continue my studies at the PhD level. It was him that made me see the beauty of syntax and to learn the social responsibility of a scholar. I also appreciate Hyunkee Ahn, Soondo Baek, HyunSook Ko, Jin Hwa Lee, Young Shik Lee, Sun Hee Park, and Sang-Keun Shin for their emotional support and encouragement.

My gratitude also goes to my participants who used their precious time to take part in the experimental studies in this dissertation. I am indebted to the people who helped me with data collection: Leilani Au, Jeffrey Bock, Jihyun Kim, Rakhun Kim, Ju Young Min, Alison Onishi, Yangon Rah, and Angie Solomon. I also would like to thank my research assistants, Les Black, Chaelyn Kamikawa, and Kristen Urada. They all generously agreed to help me with transcribing

English utterances from my participants and coding the accuracy of these transcriptions. In the process of preparing for and doing my defense, I got big help from Justin Kanda and Sunhee Kim Fujii. I am so grateful for all their support and help.

Parts of this work have been presented at BUCLD 2019, the NINJAL-UHM Workshop 2019, and the Language Acquisition Reading Group at the University of Hawai‘i. I have benefited from audience feedback in ways that are evident in this dissertation. I especially thank Dustin Crowther, Holger Hopp, Kyle Johnson, Heejeong Ko, and Jeff Lidz. I was also lucky to be able to discuss my dissertation project with Micha Elsner, Jefferey Runner, Rex Sprouse, and Kristen Syrett.

I am thankful for the financial support for this dissertation project that I got from the University of Hawai‘i (the College of Languages, Linguistics & Literature Excellence in Doctoral Dissertation Research Award, the Dai Ho Chun Fund, the Elizabeth Carr-Holmes Scholarship, and the Ruth Crymes Scholarship) and from the Fulbright Commission of Korea and the United States (IIE Grantee ID: 15151065).

It was very fortunate that I could discuss my dissertation projects with my friends at various stages of my dissertation, which I am so grateful for. Tons of thanks to Amber Camp, Hyunwoo Kim, Akari Ohba, Yangon Rah, and Fred Zenker for sharing their knowledge and wisdom with me! I also sincerely appreciate many colleagues, seniors, and friends who supported me all the way through this journey. Thank you, Bradlee-Scott Aceret, Patrick Agullana, Jenny Argueta, Les Black, Ivan Bondoc, Adam Bramlett, Yu-Tzu Chang, Sowon Cheon, Ann ‘Diane’ Choe, Peter Chong, Eunjin Chun, Bonnie Fox, Kathleen Griffin, Sanghyun Han, Eun Hye Hong, Ruadhán Hughes, Hyunji Hwang, Holly Ishiki, Hwayoung Jang, Hye Young Jung, Chaelyn Kamikawa, Donghwa Kang, Thomas Kettig, Jihyun Kim, Kitaek Kim, Ohyoung Kim, Rakhun Kim, Sangki Kim, Yeonju Kim, Elaine Lau, Eunsun Lee, Heejung Lee, Kyung Rok Lee, Victoria Lee, Yuri Lee, Yu-Han Lin, Wenyi Ling, Ju Young Min, Mari Miyao, Bora Nam, Alison Onishi, Huy Phung, Eunseok Ro, Agnese Scaturro, Anna Seo, Gyu-Ho Shin, George Smith, Hye Young Smith, Amy Song, Katherine Strong, Aya Takeda, Justin Tanaka, Nozomi Tanaka, Denis Melik Tangiyev, Kristen Urada, Jue Wang, Mayuko Yusa, Hong Ri Zenker, Jing ‘Crystal’ Zhong, and Yanxin ‘Alice’ Zhu.

I would like to express my love and deepest gratitude to my parents. My father, Chang Dong Hwang, has never said “no” to me in my life, and he was there for me all the times I

needed him, always rooting for me. My mother, Jong Sun Kim, showed me how to be thankful for all the things that I should never take for granted and how to express my thanks to others. Big thanks to my sister, Haeji Hwang, who is the best gift that I ever got in my life for free.

Above all, I thank God for his showers of blessings throughout my PhD study. I sincerely hope to be a scholar who is willing to share what I have learned with others in need, which is something that I have long prayed for.

ABSTRACT

This dissertation probes the L1 acquisition, L2 acquisition, and L2 processing of contrasts between two seemingly similar phenomena in English. The first, VP-Ellipsis (VPE), involves the deletion of an entire verb phrase (e.g., *Sara made pizza and Kelly did ~~make pizza~~ too*); the second, Gapping, involves a verb gap (e.g., *Sara made pizza and Kelly __ pasta*). One such contrast is that whereas VPE is grammatical both in conjunct clauses and in adjunct clauses (e.g., *Sara made pizza {and Kelly did too/because Kelly did}*), Gapping is grammatical only in conjunct clauses (e.g., *Sara made pizza {and Kelly pasta/*because Kelly pasta}*). Another contrast is that whereas Gapping (e.g., *Mom hugged the boy at home and Dad in the park*) allows the noun phrase following the conjunction to be interpreted as either the subject (e.g., 'hugger') or object (e.g., 'huggee') of the gapped verb, VPE (e.g., *Mom hugged the boy at home and Dad did too*) permits only a subject reading. Importantly, these grammaticality and interpretation contrasts constitute learnability challenges for L1-English children and L1-Korean L2ers of English alike: For neither group can input alone lead to implicit knowledge of the impossibility of both Gapping in adjunct clauses and the object reading for VPE; for L1-Korean L2ers, moreover, implicit knowledge that VPE in adjunct clauses is possible and that the object reading for VPE is impossible cannot come from their L1 grammar or from their classroom instruction, either.

Study 1 is a corpus-based study examining how (in)frequent VPE and Gapping are in the input to L1-English children and the input to L1-Korean L2ers of English. The input corpora to each of these groups revealed hardly any instances of VPE in adjunct clauses or of Gapping at all, which suggests that input alone cannot derive the two contrasts at issue. Two acquisition studies tested L1-English children (n = 24–33) and (early, n = 27; late, n = 30) L1-Korean L2ers of English for knowledge of these contrasts between VPE and Gapping: the grammaticality contrast via an acceptability judgment task (Study 2) and the interpretation contrast via a picture-sentence matching task (Study 3). The results showed that (a) the L1 children know the grammaticality contrast as early as age 5;11 and the interpretation contrast as early as age 5;6, and (b) the higher-proficiency early L2ers and most of the late L2ers had also acquired both contrasts. Processing of Gapping vs. VPE by adult L1-Korean L2ers (n = 48) was investigated in Study 4 via a self-paced reading task making use of the fact that (im)plausibility is manipulable

in Gapping, but not in VPE, by changing the verb (e.g., *Bill {ordered/*drank} coffee and Jane sandwiches* vs. *Bill {ordered/drank} coffee and Jane did too*). Like the native speaker controls (n = 53), the adult L2ers exhibited implausibility effects only for Gapping, thereby indicating that they can retrieve verb information at the gap site in real-time processing.

In short, the acquisition studies provide evidence that L1 children and L1-Korean L2ers can overcome the learnability problems involved in the grammaticality and interpretation contrasts between VPE and Gapping in English, and the self-paced reading study demonstrates that adult L1-Korean L2ers can process English Gapping sentences in a target-like manner.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

This dissertation employs Yale Romanization to transcribe non-English sentences. The following abbreviations are used to label linguistic terms.

ACC	Accusative case marker
AD	Adverbializer
ASP	Aspect marker
COMP	Complementizer
COP	Copula
DECL	Declarative sentence-type suffix
EX	Exclamatory suffix
GEN	Genitive particle
GER	Gerundive suffix
NOM	Nominative case marker
POL	Politeness suffix
Q	Question marker
TOP	Topic marker

CHAPTER I
INTRODUCTION

Linguists have long proposed that there exist anaphoric processes that are derived through deletion of a syntactic constituent under identity with an antecedent rather than through movement (e.g., Chomsky, 1995; Fiengo & May, 1994; Hankamer & Sag, 1976; Merchant, 2001; Sag, 1976). VP-Ellipsis (VPE) is one such construction. In VPE, a VP constituent is elided, hence not spelled out, as illustrated in (1).

(1) Mom hugged the boy at home and [Dad did [e] too].

The other construction of interest in this dissertation is Gapping (Ross, 1970), which is syntactically distinct from VPE. It is thought that this phenomenon results from Across-the-Board movement of a verb (e.g., Johnson, 2000, 2006, 2009), occasionally with another neighboring element, and that whatever is moved in one or more conjuncts of a coordinate structure becomes unpronounced, as in (2).

(2) Mom hugged the boy at home and [Dad [e] in the park[a]] / and [[e] Dad in the park[b]].

Because of the different syntactic operations involved, VPE and Gapping present an interesting contrast in terms of their grammaticality in adjunct clauses. While VPE can occur in a conjunct clause (e.g., (3a)) or an adjunct clause (e.g., (3b)), Gapping can occur only in a conjunct clause (e.g., (3c)) and not in an adjunct clause (e.g., (3d)) (Jackendoff, 1971; Lobeck, 1995).

(3) a. VPE in a conjunct clause: Sara made pizza [and Kelly did [e] too].
b. VPE in an adjunct clause: Sara made pizza [because Kelly did [e]].
c. Gapping in a conjunct clause: Sara made pizza [and Kelly [e] pasta].
d. Gapping in an adjunct clause: *Sara made pizza [because Kelly [e] pasta].

(paradigm adapted from Schwartz, 1999, p. 638, (6a)–(6d))

VPE and Gapping also contrast when it comes to interpretive ambiguity: Whereas Gapping sentences allow the argument following the conjunction (e.g., *and*) to be interpreted as either the subject or object of the gapped verb, as shown in (4a)–(4b), VPE sentences permit only a subject reading, as shown in (5a)–(5b).

(4) Mom hugged the boy at home and [Dad [e] in the park[a]] / and [[e] Dad in the park[b]].
   a. Subject reading: 'Mom hugged the boy at home, and Dad (hugged the boy) in the park.'
   b. Object reading: 'Mom hugged the boy at home, and (Mom hugged) Dad in the park.'

(5) Mom hugged the boy at home and Dad did [e] too.
   a. Subject reading: 'Mom hugged the boy at home, and Dad hugged the boy at home.'
   b. Object reading: *'Mom hugged the boy at home, and Mom hugged Dad at home.'

While there have been a few first language (L1) and second language (L2) acquisition studies on VPE and Gapping, there has been no such research as of yet on the grammaticality or interpretation contrasts between the two. Nor has the population of young ('early') L2 learners (L2ers) been the focus of any previous research on either phenomenon. This dissertation attempts to address these gaps in the literature by testing L1-English-acquiring children as well as early and late L1-Korean L2ers of English for knowledge of the above contrasts between VPE and Gapping. I expect this dissertation to expand our understanding of these understudied phenomena in L1 and L2 acquisition, allowing us both to make comparisons across L1 acquisition, early L2 acquisition, and late L2 acquisition and to draw theoretical conclusions.

What makes these contrasts between VPE and Gapping particularly interesting is that they constitute learnability problems both in the context of L1 acquisition and in the context of key concern to us: the L2 acquisition of English by L1-Korean speakers (see Crain, 1991; Schwartz & Sprouse, 2000, 2013). In neither of these contexts can the input alone account for the acquisition of implicit knowledge that both Gapping in adjunct clauses and the object reading for VPE are impossible. Nor can implicit knowledge of the impossibility of Gapping in adjunct clauses be acquired via analogy with either Gapping in conjunct clauses or VPE in adjunct clauses, since Gapping in adjunct clauses is impossible but Gapping in conjunct clauses and VPE in adjunct clauses are both possible. Likewise, implicit knowledge of the impossibility of VPE

with an object reading cannot be derived from analogy with either VPE with a subject reading or Gapping with an object reading because VPE with an object reading is impossible but VPE with a subject reading and Gapping with an object reading are both possible. For L1-Korean L2ers, moreover, implicit knowledge that VPE in adjunct clauses is possible and that the object reading for VPE is impossible cannot come from their L1 grammar or from their classroom instruction (see §2.4).

From a processing perspective, Gapping also provides a useful probe to investigate whether and, if so, when a missing verbal element is processed in real time. However, Gapping is still understudied in native language processing (for Dutch, see Kaan, Overfelt, Tromp, & Wijnen, 2013; for English, see Carlson, 2001, 2002; Carlson, Dickey, & Kennedy, 2005; Hoeks, Redeker, & Hendriks, 2009; Kaan, Wijnen, & Swaab, 2004; N. Kim, Carlson, Dickey, & Yoshida, 2020; for German, see Claus, 2015; Hofmann, 2006; Streb, Hennighausen, & Rösler, 2004) and has never been studied in L2 processing. To address this gap in the research, one of the experimental studies in this dissertation investigates adult L2 processing of Gapping vs. VPE using a self-paced reading task. This task makes use of the fact that the plausibility of a sentence can be manipulated by changing the verb in the case of Gapping (e.g., (6a) vs. *(6b)) but not in the case of VPE (e.g., (6c) vs. (6d)) (see Kaan et al., 2004).

(6) a. Bill ordered coffee and tea at the cafe, and Jane [e] sandwiches and cake at the bakery.
b. *Bill drank coffee and tea at the cafe, and Jane [e] sandwiches and cake at the bakery.
c. Bill ordered coffee and tea at the cafe, and Jane did [e] too.
d. Bill drank coffee and tea at the cafe, and Jane did [e] too.

Given that previous research on the L2 processing of filler–gap dependencies has focused primarily on *wh*-questions and relative clauses, the current online Gapping study will contribute new information to our understanding of the processing capabilities of adult L2ers.

The dissertation is organized as follows: Chapter 2 lays out the key syntactic properties of VPE (§2.1) and Gapping (§2.2) with a focus on the differences between English and Korean. It also presents the target grammaticality and interpretation contrasts between VPE and Gapping in English (§2.3), and discusses why these contrasts constitute learnability problems for L1-English-acquiring children and L1-Korean L2ers of English (§2.4). Chapter 3 reviews

previous research on the L1 acquisition of VPE (§3.1), the L2 acquisition of VPE (§3.2) and Gapping (§3.3), and the L1 processing of Gapping (§3.4). Chapter 4 reports separate corpus-based natural language processing analyses of English input data to, respectively, L1 children and L1-Korean L2ers in Korea; they reveal that exposure to the Target Language is unable to provide either population with direct information about the contrasts between VPE and Gapping in English. The next two chapters present the details and the results of the English acquisition experiments with both L1 children and (early and late) L1-Korean L2ers on the VPE vs. Gapping grammaticality contrast (Chapter 5) and the VPE vs. Gapping interpretation contrast (Chapter 6). Chapter 7 provides the methods and results of the real-time L2 processing study with L1-Korean adults on English Gapping. Chapter 8 summarizes the results of the four studies in §8.1, ties together the results of the two acquisition studies in §8.2, discusses theoretical implications in §8.3, and finally conclude this dissertation in §8.4.

CHAPTER II
LINGUISTIC BACKGROUND:
VP-ELLIPSIS AND GAPPING IN ENGLISH AND KOREAN

This chapter provides an overview of syntactic analyses of VP-Ellipsis (VPE) and Gapping in English and Korean. In §2.1, I first overview how English VPE is analyzed. Next, I discuss three Korean constructions (Argument Ellipsis, *Kulay* 'Do So' Anaphora, and Pseudo-VP-Ellipsis) that some have argued to be close analogues of English VPE and explain why none of them can be treated as equivalent to English VPE. In §2.2, I turn to syntactic analyses of Gapping in English and Korean. Section 2.3 presents the target phenomena in this dissertation—the grammaticality and interpretation contrasts between VPE and Gapping in English, and §2.4 discusses why they constitute learnability problems both for L1-English acquiring children and for L1-Korean L2ers of English. Section 2.5 provides a summary of the chapter.

2.1 VP-Ellipsis in English and Korean

VPE involves a distinct type of dependency relation between an elided VP and its antecedent, as exemplified in (1).

(1) Sara made pizza and Kelly did [e] too.

This dependency is typically treated in terms of syntactic relations (Chomsky, 1995; Chomsky & Lasnik, 1993; Fiengo & May, 1994; Hankamer & Sag, 1976; Hestvik, 1995; Merchant, 2001; Rouveret, 2012; Sag, 1976; Sag & Hankamer, 1984), although it has been also approached from

a semantic perspective (Dalrymple, Stuart, & Fernando, 1991; Hardt, 1993),[1] a discourse perspective (Kehler, 2000, 2002), and a processing perspective (Arregui, Clifton, Frazier, & Moulton, 2006).[2] The current study views VPE as a structure in which one VP undergoes deletion under the condition of identity with another VP in the given syntactic or discourse context.[3]

For example, the English sentence (1) involves VP-deletion as diagramed in (2), where the boxed phrase is elided. The VP-deletion operation is licensed by the T head (e.g., Merchant, 2001). When a VP is elided in English, the stranded Tense feature under T needs to be supported by various types of verbal elements, such as *do*, copula *be*, auxiliaries (*be*, *have*), and modals

---

[1] Semantic accounts claim that VPE involves a purely semantic identity relation between the ellipsis site and its antecedent. According to Dalrymple et al. (1991), the semantic identity is established over the meanings of predicates. To illustrate, the semantic interpretation of the antecedent clause in (i) corresponds to (ii). It can be represented as property *P* being predicated of *dan* by the lambda term as in (iii). If a value, *dan*, is substituted for the free variable, *x*, property *P* can be represented as a function in (iv). When this function applies to the ellipsis clause in (i), a parallel interpretation 'Dan likes golf, and George likes golf too' is obtained.

(i) Dan likes golf, and George does too.

(Dalrymple et al., 1991; p. 5, (4))

(ii) *like* (*dan, golf*)

(iii) $P \rightarrow \lambda x.like$ (*dan, golf*)

(iv) $P \rightarrow \lambda x.like$ (*x, golf*)

[2] Arregui et al. (2006) propose the "VP recycling hypothesis," which states that the grammatical resolution of an elided VP requires the presence of a syntactically identical antecedent. However, when such a syntactically parallel antecedent is absent, as in (v), the processor may recycle material available from the syntactically unparallel antecedent and create the appropriate VP, as in (vi), as a performance repair strategy.

(v) *The dessert was praised by the customer and the critic did.

(Arregui et al., 2006, p. 241, (16b))

(vi) The dessert was praised by the customer and the critic did [praise the dessert].

[3] Note that morphological identity is not strictly required for VPE, as shown in (vii).

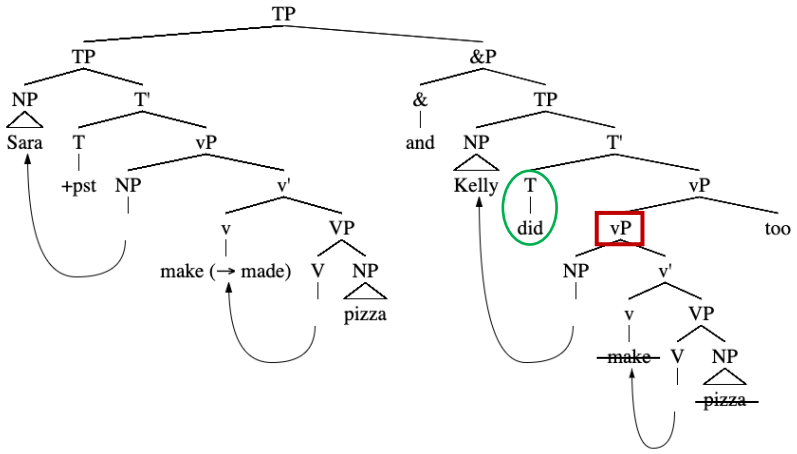
(vii) a. Mike revises his work, and Jennifer should ~~revise her work~~ as well.
b. John plays football, but Tom hasn't ~~played football~~ recently.

(adapted from Al-Thubaiti, 2019, p. 234, Table 2, (1)–(2))

According to Lasnik (1995), (viia) and (viib) are grammatical because lexical verbs (e.g., *revise* in (viia) and *play* in (viib)) are inflected at the level of syntax where the VP is elided, which makes it possible for morphological identity to be ignored. I return to this issue in §3.2.

(e.g., *can*, *could*). In particular, *do*-support is employed when the antecedent clause is in the simple present or simple past. This operation is indicated with a circle in (2).

(2)



(adapted from the analysis of Sag, 1976, p. 18, (1.2.4))

The consensus for Korean is that there is no VPE (e.g., H.-D. Ahn, 2018; Goldberg, 2005; J.-S. Kim, 2006; M.-K. Park, 1997). Some may argue that Korean Argument Ellipsis (Cho, 2001; Cole, 1987; Huang, 1984, 1991; Otani & Whitman, 1991), *Kulay* 'Do So' Anaphora, and Pseudo-VP-Ellipsis (J.-S. Kim, 1997) are close analogues of English VPE. I discuss each of these in detail to build the case that none can be considered a direct counterpart to English VPE.

Before doing this, however, I introduce one construction in Mandarin, the first East Asian language investigated for VPE-like phenomena. Huang (1984, 1988, 1991) claimed that VPE in Mandarin (and other East Asian languages) is realized, as in (3), by repeating a lexical verb from the first clause without pronouncing its object.

(3) John diudiao-le ta-de xin Bill ye diudiao-le [e].
John discard-ASP he-GEN letter Bill also discard-ASP
'John discarded his letter and Bill did too.'

Korean has a construction that seemingly works the same way as Mandarin (3) (Cho, 2001; Cole, 1987; Huang, 1984, 1991; Otani & Whitman, 1991), as shown in (4).

(4) John-i ku-uy phyenci-lul pely-ess-ko Bill-to [e] pely-ess-e-yo.
John-NOM he-GEN letter-ACC discard-PST-and Bill-also discard-PST-DECL-POL
'John discarded his letter and Bill did too.'

However, Huang's (1984, 1988, 1991) claim about VPE in East Asian languages was challenged by Goldberg (2005), Hoji (1998), H.-J. G. Li (1998), and Xu (2003), among others. According to them, the construction shown in (3) and (4) is not analogous to English VPE. Compare the English sentence in (5) and the Korean sentence in (6), focusing on the interpretation of the adverbial in the ellipsis clause.

(5) Tom ran fast and Kyle did [e] too.

(6) Tom-i ppalli ttwi-ess-ko Kyle-to [e] ttwi-ess-e-yo.
Tom-NOM fast run-PST-and Kyle-also run-PST-DECL-POL
'Tom ran fast and Kyle did too.'
'Tom ran fast and Kyle ran (at some unspecified speed).'

In English VPE, a VP containing an adverbial in the antecedent clause is thought to be represented identically in the subsequent ellipsis clause. The only possible interpretation of (5) is therefore that *Tom* and *Kyle* both ran fast. However, the Korean sentence in (6) allows an alternative interpretation in which the adverbial only describes the action in the first clause, i.e., 'Tom ran fast, and Kyle ran (at some unspecified speed).'

Moreover, the Korean sentence containing negation in (7) has a different interpretation with regard to the adverb than the English VPE equivalent in (8) does.

(7) Tom-un ppalli ttwi-ess-ciman Kyle-un [e] ttwi-ci anh-ass-e-yo.
Tom-TOP fast run-PST-but Kyle-TOP run-COMP NEG-PST-DECL-POL
'Tom ran fast but Kyle didn't run.'

(8) Tom ran fast but Kyle didn't [e].

While the only interpretation for the ellipsis clause in (8) is 'Kyle ran, but not fast,' such an interpretation cannot be obtained from the Korean sentence (7). The only possible interpretation of the ellipsis clause in (7) is 'Kyle did not run at all.'

In addition, the same Korean construction is acceptable even if the verbs are different between the two clauses, as in (9).[4]

(9) John-i phyenci-lul pely-ess-ko Bill-i [e] cwu-wess-e-yo.
John-NOM letter-ACC discard-PST-and Bill-NOM pick up-PST-DECL-POL
'John discarded a letter and Bill picked (it) up.'

To further complicate the problem, this type of Korean construction permits other readings aside from the so-called 'sloppy' and 'strict' readings licensed by English VPE for a pronoun elided in the ellipsis clause. For example, the English sentence (10) only allows the readings in (11) (Foley, Núñez del Prado, Barbier, & Lust, 2003; Matsuo, 2007; Thornton & Wexler, 1999).

(10) John discarded his letter and Bill did [e] too.

(11) Interpretations

Sloppy:

a. $John_i$ discarded $John_i$'s letter, and $Bill_j$ discarded $Bill_j$'s letter.

Strict:

b. $John_i$ discarded $John_i$'s letter, and $Bill_j$ discarded $John_i$'s letter.

c. $John_i$ discarded $Bill_j$'s letter, and $Bill_j$ discarded $Bill_j$'s letter.

d. $John_i$ discarded somebody $else_k$'s letter, and $Bill_j$ discarded somebody $else_k$'s letter.

By contrast, the Korean sentence in (4) permits other readings as well, such as '$John_i$ discarded $John_i$'s letter and $Bill_j$ discarded a letter/letters other than $John_i$'s or $Bill_j$'s.' (Hoji, 1998).

Based on the evidence above, I conclude that the Korean phenomena exemplified in (4), (6), and (7) are not equivalent to English VPE. Goldberg (2005), Hoji (1998), and S. Kim (1999)

[4] My thanks to Bonnie D. Schwartz (personal communication, 9 May 2018) for leading me to see this.

analyze sentences like (4) as Argument Ellipsis (AE)[5] in which the empty category is an NP whose sole content is its head N (e.g., *letter* in (4)). According to Hoji (p. 142), then, the reason why this construction (like English VPE) allows both strict and sloppy readings is that the null NP is behaving like either a definite or indefinite NP. In the definite use, the missing NP denotes the NP antecedent, resulting in the strict interpretation. For example, the empty position in (4) to be recovered in the ellipsis clause would be 'the letter (John's letter).' The indefinite use of the deleted NP would give rise to the other interpretations where something equivalent to 'letter' is recovered, including the sloppy interpretation. In sum, Hoji maintains that the dropped argument in AE is an NP containing a bare nominal and that its interpretation is determined by the context. However, Hoji's analysis does not provide an explanation for the ambiguity contrast between (6) and (7) with respect to adverbs. I leave this question open for future work.

An alternative candidate for VPE in Korean is the construction containing *kule(h)* 'so,' which is interpreted in the same manner as English VPE concerning adverbials (e.g., (5), (8)) and pronouns (e.g., (10)). As shown in (12), the vP in the second clause, whose meaning is identical to the meaning of the vP in the first clause, is realized as *kulay* (a contraction of *kule-hay*) 'do so.'

(12) John-i [$_{vP}$ phyenci-lul pely]-ess-ko Bill-to [$_{vP}$ kulay(=kule-hay)]-ss-e-yo.
John-NOM letter-ACC discard-PST-and Bill-also do.so(=so-do)-PST-DECL-POL
'John discarded a letter, and Bill did so too.'

It should be pointed out that *kule(h)* 'so' can also take the place of other syntactic constituents, such as AdvP and CP, as shown in (13) and (14), respectively.

(13) Cheli-ka ppalli ttwi-ess-ko Yengi-to kuleh-key ttwi-ess-e-yo.
Cheli-NOM fast run-PST-and Yengi-also so-AD run-PST-DECL-POL
'Cheli ran fast, and Yengi ran so, too.'

(adapted from M.-K. Park, 2015, p. 694, (3a))

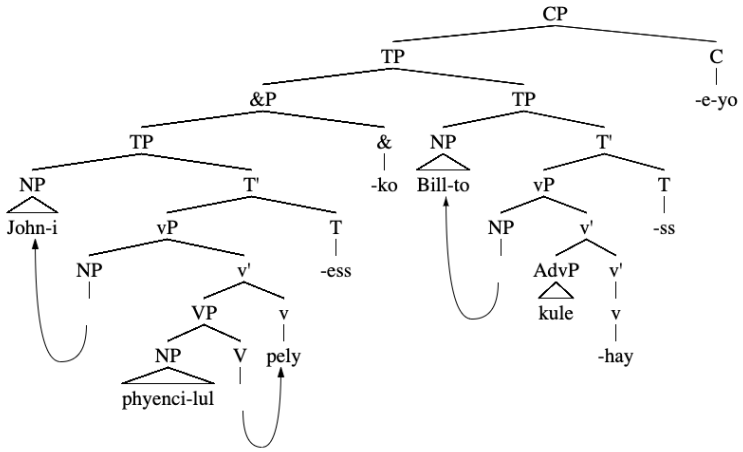
[5] Hoji (1998) limits his discussion to sentences in which an object has been dropped (e.g., (4)), which he calls "Null Object Constructions" (NOCs). However, Hoji's observations regarding NOCs are also true of sentences with null subjects, as will be exemplified in (25) (see also Goldberg, 2005; S. Kim, 1999). In this dissertation, I therefore use the term "Argument Ellipsis" instead.

(14) A: Emci-nun cengmal yeyppu-kwuna!
Emci-TOP really beautiful-EX
'Emci is really cute!'
B: Ne-to kuleh-key (= Emci-ka cengmal yeypputa-ko) sayngkakha-ni?
You-also so-AD (= Emci-NOM really beautiful-COMP) think-Q
'Do you think so, too?'

(adapted from M.-K. Park, 2015, p. 694, (3b))

Given the flexibility in the use of *kule(h)* 'so,' as in (13) and (14), the phenomenon at issue cannot be treated on a par with VPE but instead is more akin to *Do So* Anaphora in English (M.-K. Park, 2015). Based on earlier work by Stroik (2001) and Houser (2010), M.-K. Park proposed that the *kulay (kule-hay)* 'do so' in (12) has no bearing on the internal structure of its antecedent vP. Instead, *kulay (kule-hay)* forms a vP from the beginning of the derivation with *hay* 'do' as the head and *kule(h)* 'so' as an "obligatory adverb" (Houser, 2010, p. 36), as shown in (15).

(15)



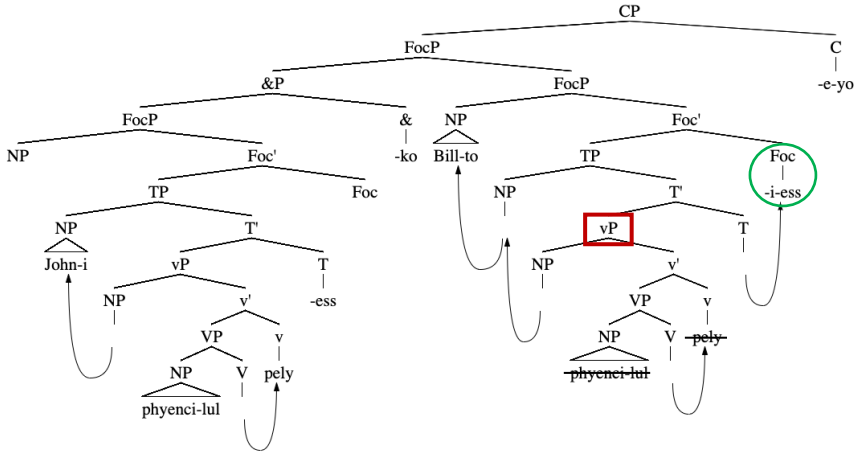
(adapted from the analysis of M.-K. Park, 2015)

The last candidate I discuss for VPE in Korean is the construction containing a focused constituent followed by *-to* 'also' and the copula, which is termed "Pseudo-VP-Ellipsis" (Pseudo-VPE; J.-S. Kim, 1997). An example of this in Korean is provided in (16), wherein the subject *Bill* is accompanied by a focus particle *-to* 'also' and the copula *-i*.

(16) John-i phyenci-lul pely-ess-ko Bill-to [e] -i-ess-e-yo.
John-NOM letter-ACC discard-PST-and Bill-also -COP-PST-DECL-POL
'John discarded a letter, and Bill did too.'

This construction involves (a) movement of the focused constituent to Spec,FocP, which is licensed by the Focus head *-i* 'be,' and (b)VP-deletion (J.-S. Kim, 1997), as represented in (17).[6]

(17)



(adapted from the analysis of J.-S. Kim, 1997)

Crucially, Pseudo-VPE generates the same interpretations as English VPE does with regard to a VP adverbial (e.g., (5), (8)) and a pronoun (e.g., (10)). However, Pseudo-VPE employs the copula *-i* 'be' no matter what type of verb is used in the antecedent clause. If the verb *ha-* 'do' is used instead, this results in ungrammaticality, as shown in (18). That is, *ha*-support cannot license the ellipsis of a VP in Korean (J. Kim, 2012).

(18) *John-i phyenci-lul pely-ess-ko Bill-to [e] hay-ss-e-yo.
John-NOM letter-ACC discard-PST-and Bill-also do-PST-DECL-POL

[6] William O'Grady (personal communication, 29 April 2020) sees J.-S. Kim's (1997) analysis as problematic because the copula *-i* 'be' in Korean cannot take a VP complement. I leave this issue for further research.

The other particularly important difference has to do with the fact that the copula need not be marked for tense, as in (19).

(19) John-i phyenci-lul pely-ess-ko Bill-to [e] -i-ey-yo.
John-NOM letter-ACC discard-PST-and Bill-also -COP-DECL-POL
'John discarded a letter, and Bill did too.'

Based on the insufficiency of a VPE analysis of the Korean data, this dissertation assumes that AE, *Kulay* 'Do So' Anaphora, and Pseudo-VPE are not analogous to English VPE. Importantly, I note here further syntactic differences between VPE, on the one hand, and the three other constructions discussed, on the other. One concerns their grammaticality in adjunct clauses. Specifically, VPE is grammatical not only in conjunct clauses (e.g., (1), (10)) but also in adjunct clauses, as in (20) (Jackendoff, 1971; Lobeck, 1995).

(20) John discarded a letter because Bill did [e].

By contrast, AE, *Kulay* 'Do So' Anaphora, and Pseudo-VPE in Korean are all disallowed in adjunct clauses, as shown in (21)–(23), except for one case: As (21) demonstrates, AE is acceptable when an overt argument in the ellipsis clause has the nominative case marker, but not when it is marked with *-to* 'also.'

(21) <u>AE</u>
Bill-i/*-to [e] pely-ess-ki-ttaymwuney, John-i phyenci-lul pely-ess-e-yo.
Bill-NOM/-also discard-PST-NM-because John-NOM letter-ACC discard-PST-DECL-POL
'John discarded a letter because Bill_{NOM} did.'

(22) <u>*Kulay* 'Do So' Anaphora</u>
*Bill-i/-to kulay-ss-ki-ttaymwuney, John-i phyenci-lul pely-ess-e-yo.
Bill-NOM/-also do.so-PST-NM-because John-NOM letter-ACC discard-PST-DECL-POL

(23) Pseudo-VPE

*Bill-i/-to	[e]	-i-ess-ki-ttaymwuney,	John-i	phyenci-lul	pely-ess-e-yo.
Bill-NOM/-also		-COP-PST-NM-because	John-NOM	letter-ACC	discard-PST-DECL-POL

One may be tempted to attribute the ungrammaticality of *Kulay* 'Do So' anaphora and Pseudo-VPE in adjunct clauses to the order of the two clauses. That is, because the adjunct clause involving anaphora or ellipsis precedes the antecedent clause, as in (22)–(23), it gives rise to an inappropriate context for the parser to process the clause containing anaphora/ellipsis. This account, however, is implausible for two reasons. First, such an order does not necessarily result in unacceptability in other languages. For instance, the sentence *Because Bill did, John discarded a letter* is acceptable in English. Second, the ungrammaticality of Pseudo-VPE does not change even if a felicitous context is added before that adjunct clause. For example, even with the context in (24a), the second sentence with an adjunct clause containing Pseudo-VPE is still unacceptable, as in (24b); it cannot mean 'Ann cried because Bill also discarded Ann's letter (as Tom did)' or 'Ann cried because Bill did.' However, note that *Kulay* 'Do So' Anaphora is acceptable in adjunct clauses when it follows its antecedent that is in a separate sentence or main clause, as in (24c), which can only mean 'Ann cried because Bill also discarded Ann's letter (as Tom did).'[7]

(24) a.

Tom-i	Ann-uy	phyenci-lul	pely-ess-e-yo.
Tom-NOM	Ann-GEN	letter-ACC	discard-PST-DECL-POL

'Tom discarded Ann's letter.'

b. Pseudo-VPE

*Bill-to	[e]	-i-ess-ki-ttaymwuney,	Ann-i	wul-ess-e-yo.
Bill-also		-COP-PST-NM-because	Ann-NOM	cry-PST-DECL-POL

c. *Kulay* 'Do So' Anaphora

Bill-to	kulay-ss-ki-ttaymwuney,	Ann-i	wul-ess-e-yo.
Bill-also	do.so-PST-NM-because	Ann-NOM	cry-PST-DECL-POL

[7] I am grateful to William O'Grady (personal communication, 6 December 2018) for bringing this to my attention.

Another difference between VPE and the three Korean phenomena under discussion is that the Korean constructions allow arguments other than the subject to come before *-to* 'also' in the non-antecedent clause, as shown in (25)–(27).

(25) AE

John-i	phyenci-lul	pely-ess-ko	khatu-to	[e]	pely-ess-e-yo.
John-NOM	letter-ACC	discard-PST-and	card-also		discard-PST-DECL-POL

'John discarded a letter, and (he) discarded a card too.'

(26) *Kulay* 'Do So' Anaphora

John-i	phyenci-lul	pely-ess-ko	khatu-to	kulay-ss-e-yo.
John-NOM	letter-ACC	discard-PST-and	card-also	do.so-PST-DECL-POL

'John discarded a letter, and (he discarded) a card too.'

(27) Pseudo-VPE

John-i	phyenci-lul	pely-ess-ko	khatu-to	[e]	-i-ess-e-yo.
John-NOM	letter-ACC	discard-PST-and	card-also		-COP-PST-DECL-POL

'John discarded a letter, and (he discarded) a card too.'

As a consequence, ambiguity can arise in these three constructions. For example, in (28)–(30), the argument following the conjunction (e.g., *-ko* 'and') can have either a subject reading (SR), as in (28a), or an object reading (OR), as in (28b).

(28) AE

Emma-ka	cip-eyse	sonyen-ul	an-ass-ko	appa-to	[e]	an-ass-e-yo.
Mom-NOM	home-at	boy-ACC	hug-PST-and	Dad-also		hug-PST-DECL-POL

a. SR: 'Mom hugged the boy at home, and Dad hugged the boy at home.'

b. OR: 'Mom hugged the boy at home, and Mom hugged Dad at home.'

(29) *Kulay* 'Do So' Anaphora

Emma-ka	cip-eyse	sonyen-ul	an-ass-ko	appa-to	kulay-ss-e-yo.
Mom-NOM	home-at	boy-ACC	hug-PST-and	Dad-also	do.so-PST-DECL-POL

(30) Pseudo-VPE

Emma-ka	cip-eyse	sonyen-ul	an-ass-ko	appa-to	[e]	-i-ess-e-yo.
Mom-NOM	home-at	boy-ACC	hug-PST-and	Dad-also		-COP-PST-DECL-POL

However, English VPE sentences, as in (31), are unambiguous, permitting only SR in (31a).

(31) Mom hugged the boy at home and Dad did [e] too.

a. SR: 'Mom hugged the boy at home, and Dad hugged the boy at home.'

b. OR: *'Mom hugged the boy at home, and Mom hugged Dad at home.'

The evidence above makes it clear that AE, *Kulay* 'Do So' Anaphora, and Pseudo-VPE cannot be considered exact counterparts to English VPE. In AE, an adverbial in the antecedent clause does not need to be recovered in the ellipsis clause. AE also allows readings other than the so-called 'sloppy' and 'strict' readings for a pronoun elided in the ellipsis clause. Moreover, AE is acceptable even in the case where the verbs are different in the antecedent and ellipsis clauses. Pseudo-VPE is unlike English VPE in that it is ungrammatical in adjunct clauses. *Kulay* 'Do So' Anaphora is also disallowed in adjunct clauses except when its antecedent precedes it in a separate sentence or main clause. Furthermore, all these Korean constructions differ from English VPE in that they can be ambiguous with regard to the interpretation of the argument following the conjunction. I therefore conclude that Korean does not have an exact equivalent to VPE (see also H.-D. Ahn, 2018; Goldberg, 2005; J.-S. Kim, 2006; M.-K. Park, 1997).

2.2 Gapping in English and Korean

Another construction that exhibits superficial similarity to VPE is what Ross (1970) called "Gapping." Gapping suppresses the verb (occasionally with another neighboring element) in one or more clauses of a coordinate structure under conditions of identity with the verb (and

the neighboring element) in the other clause (Johnson, 2000, 2006, 2009). In English, it applies forward, the result of which is that the verb in the second clause is unpronounced, as in (32).

(32) Sara made pizza and Kelly [e] pasta.

A similar phenomenon is observed in Korean, as shown in (33).

(33) Kelly-ka/-nun phasutha-lul [e] (kuliko) Sara-ka/-nun phica-lul mantul-ess-e-yo.
Kelly-NOM/-TOP pasta-ACC (and) Sara-NOM/-TOP pizza-ACC make-PST-DECL-POL
'Sara made pizza and Kelly pasta.'

However, Gapping is realized somewhat differently in Korean than in English. First, the suppression of the verb is in the first clause in Korean. Second, the first argument in the two clauses in Korean Gapping can be marked with either a case marker (e.g., Accusative: *-ul/-lul*; Nominative: *-i/-ka*) or a topic marker (i.e., *-un/-nun*).

The first difference between English and Korean Gapping can be accounted for by O'Grady's (1999) Constraint on Gapping Direction. Building on Ross (1970) and Johannessen (1996), he proposed that head-complement order predicts "the impossibility of a particular gapping direction without implying that the reverse gapping direction is permitted" (p. 143), as schematized in (34).

(34) Constraint on Gapping Direction

a. Verb-Object languages (e.g., English): *[S … Ø …] [S … V …]

b. Object-Verb languages (e.g., Korean): *[S …… V] [S …… Ø]

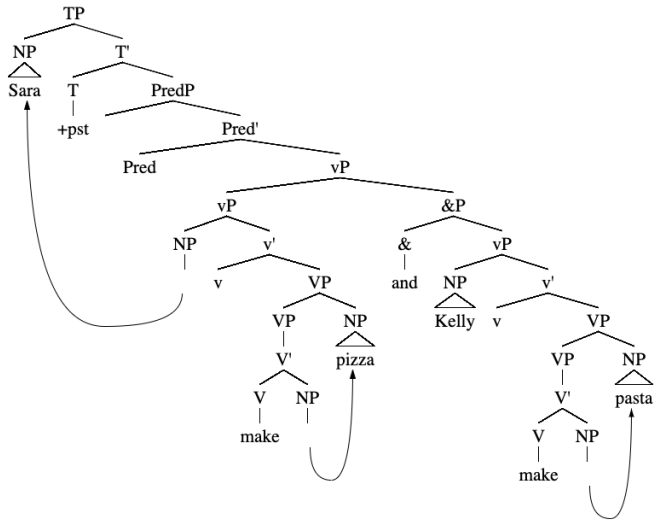
(adapted from O'Grady, 1999, p. 143, (6))

Regarding the syntactic analysis of Gapping, this dissertation follows Johnson (2000, 2006, 2009), who considers it to be an instance of Across-the-Board (ATB) movement.[8] According to this proposal, Gapping arises when coordination takes place at the vP level. The

[8] For a multiple dominance analysis, see Chung (2004); for a sideward movement analysis, see Agbayani and Zoerner (2004); and for a string deletion analysis, see Mukai (2003).

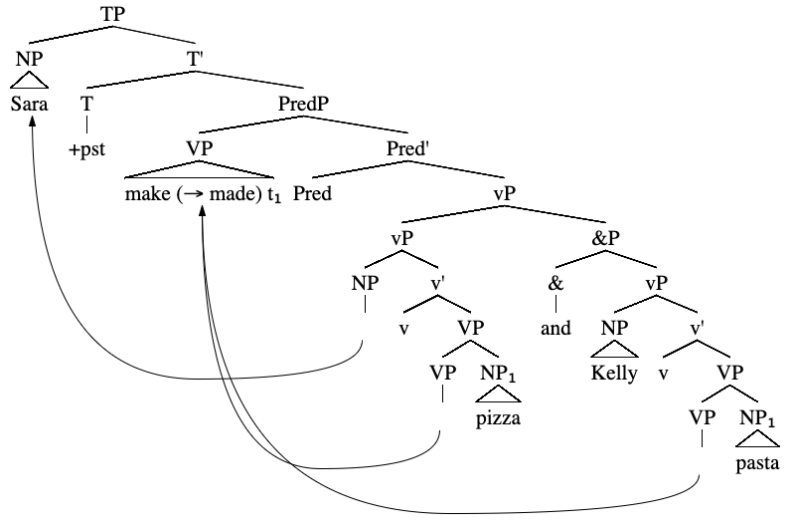
English Gapping sentence in (32), for example, can be analyzed such that the complements of *make*, namely *pizza* and *pasta*, first move out of their original positions through rightward Heavy NP Shift to receive focus, as in (35). These moved elements are termed "remnants."

(35) Step 1: Heavy NP shift in English Gapping



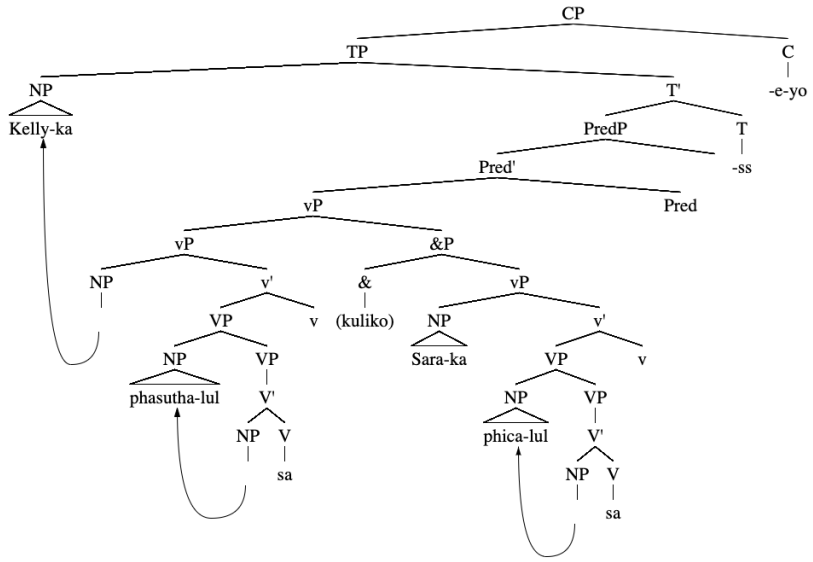
This operation feeds the ATB movement of the VPs containing its head *make* up to the Specifier of PredP (Zwart, 1997), as in (36). In this process, the ATB movement of VP is assumed to be licensed by Pred (Johnson, 2009, p. 307).

(36) Step 2: Leftward ATB movement in English Gapping

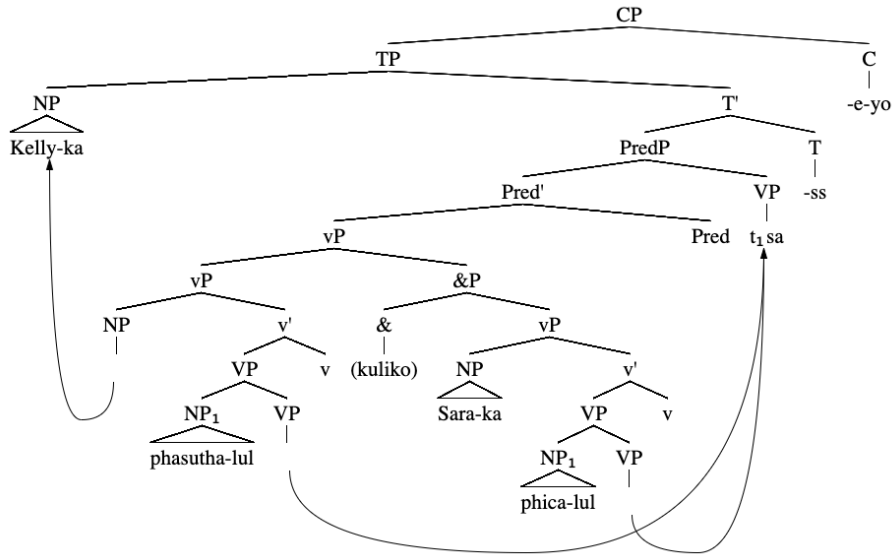


The Korean Gapping sentence in (33) can be analyzed in the same way, as shown in (37) and (38), although the direction of (non-subject) movement is different.

(37) Step 1: Heavy NP shift in Korean Gapping



(38) Step 2: Rightward ATB movement in Korean Gapping



The ATB movement account assumes that Gapping arises when vPs have been coordinated. Therefore, Gapping is possible in the conjunct clause in (32), but not in the adjunct clause in (39).

(39) *Sara made pizza because Kelly [e] pasta.

The same contrast in grammaticality holds for Korean Gapping. Korean Gapping is grammatical in conjunct clauses (Sohn, 1999), but not in adjunct clauses, as shown in (40) vs. (33).

(40) *Kelly-ka phasutha-lul [e] ttaymwuney, Sara-ka phica-lul mantul-ess-e-yo.
Kelly-NOM pasta-ACC because Sara-NOM pizza-ACC make-PST-DECL-POL
'*Sara made pizza because Kelly pasta.'

On the other hand, English Gapping sentences exhibit an interesting pattern regarding their interpretation. As shown in (41), they are ambiguous such that the argument following the coordinating conjunction (e.g., *and*) can have either an SR or an OR (see Carlson, 2001; Carlson et al., 2005; Kaan et al., 2004).

(41) Mom hugged the boy at home and Dad [e] in the park[a] / and [e] Dad in the park[b].
a. SR: 'Mom hugged the boy at home, and Dad (hugged the boy) in the park.'
b. OR: 'Mom hugged the boy at home, and (Mom hugged) Dad in the park.'

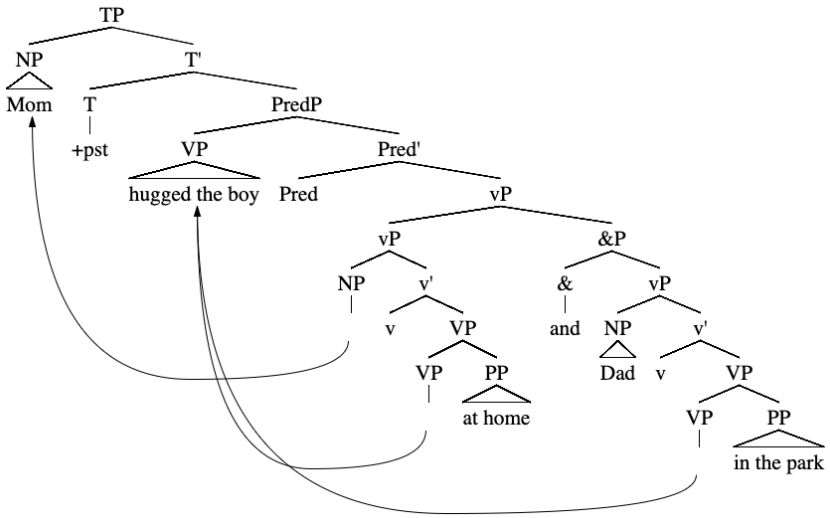
The SR in (41a) and the OR in (41b) are associated with two structurally distinct parses: SRs involve vP coordination (e.g., (42)) and ORs involve VP coordination (e.g., (43)).[9]

[9] Robert Bley-Vroman (personal communication, 28 April 2020) brought another analysis of ORs to my attention. According to Jackendoff (1971; see also Carlson, 2001), ORs involve conjunction reduction. Thus, the sentence in (41) with an OR is derived by merely deleting material from the left-hand side of the second clause, as illustrated in (viii).

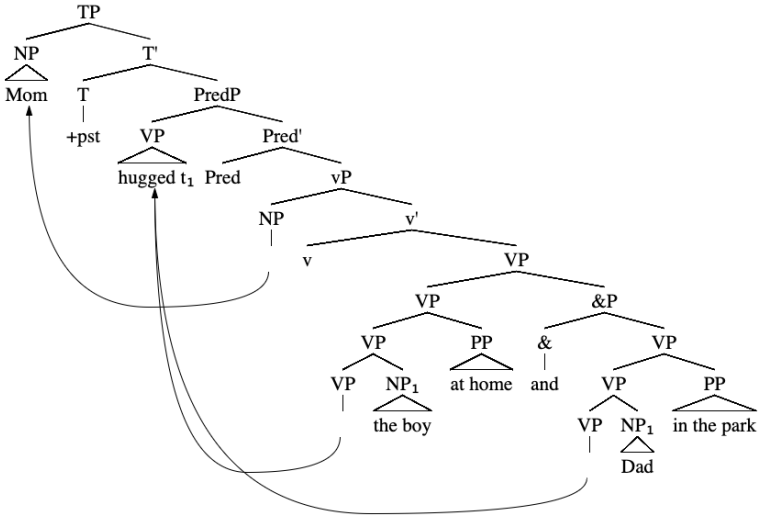
(viii)Mom hugged the boy at home and ~~Mom hugged~~ Dad in the park.

I will not delve further into which analysis of ORs is preferred because that would go beyond the scope of this dissertation.

(42) SR



(43) OR



(adapted from the analyses of Johnson, 2000, 2009 & Zoerner, 1999)

Korean Gapping sentences also allow both an SR and an OR, but the ambiguity can be resolved through the use of an accusative case marker or a nominative case marker, as demonstrated in (44) and (45).

(44) Appa-ka/nun kongwen-eyse [e] emma-ka/nun cip-eyse sonyen-**ul** an-ass-e-yo.
Dad-NOM/TOP park-in Mom-NOM/TOP home-at boy-**ACC** hug-PST-DECL-POL
SR: 'Mom hugged the boy at home, and Dad (hugged the boy) in the park.'

(45) Appa-lul/nun kongwen-eyse [e] sonyen-ul/un cip-eyse emma-**ka** an-ass-e-yo.
Dad-ACC/TOP park-in boy-ACC/TOP home-at Mom-**NOM** hug-PST-DECL-POL
OR: 'Mom hugged the boy at home, and (Mom hugged) Dad in the park.'

However, if all arguments have a topic marker and if the shared argument of the gapped and non-gapped clauses is fronted to sentence-initial position, then Korean Gapping sentences become ambiguous, albeit in a different way from English Gapping sentences. As shown in (46), the fronted argument can be interpreted as either the shared subject or the shared object of the gapped and non-gapped clauses.

(46) Emma-nun$_i$ appa-nun kongwen-eyse [e] sonyen-un cip-eyse t_i an-ass-e-yo.
Mom-TOP Dad-TOP park-in boy-TOP home-at hug-PST-DECL-POL
Shared SR: 'Mom hugged the boy at home and Mom hugged Dad in the park.'
Shared OR: 'The boy hugged Mom at home and Dad hugged Mom in the park.'

2.3 Target Phenomena in the Present Dissertation

The target phenomena in this dissertation are the grammaticality and interpretation contrasts between VPE and Gapping in English that were discussed in §2.1 and §2.2. Table 2.1 shows the grammaticality of VPE, AE, *Kulay* 'Do So' Anaphora, Pseudo-VPE, and Gapping in conjunct and adjunct clauses and Table 2.2 shows the distribution of possible interpretations of these phenomena.

Table 2.1
Grammaticality of VPE, AE, Kulay 'Do So' Anaphora, Pseudo-VPE, and Gapping

	English		Korean	
Construction	**Conjunct**	**Adjunct**	**Conjunct**	**Adjunct**
VPE	✓	✓		
AE			✓	✓[a]/*
***Kulay* 'Do So' Anaphora**			✓	✓[b]/*
Pseudo-VPE			✓	*
Gapping	✓	*	✓	*

Notes. [a]AE is grammatical in adjunct clauses only when an overt argument in the ellipsis clause has the nominative case marker (vs. *-to* 'also'). [b]*Kulay* 'Do So' Anaphora is grammatical in adjunct clauses only when its antecedent precedes it in a separate sentence or main clause.

Table 2.2
Interpretation of VPE, AE, Kulay 'Do So' Anaphora, Pseudo-VPE, and Gapping

	English		Korean	
Construction	**SR**	**OR**	**SR**	**OR**
VPE	✓	*		
AE			✓	✓
***Kulay* 'Do So' Anaphora**			✓	✓
Pseudo-VPE			✓	✓
Gapping	✓	✓	✓	✓

Critical to this dissertation is the fact that the grammaticality and interpretation contrasts between VPE and Gapping in English constitute learnability problems for both L1-English children and L1-Korean L2ers of English. This topic is discussed in detail in the next section.

2.4 Learnability Issues Involved in the Target Phenomena in the Present Dissertation

2.4.1 Grammaticality contrast between VP-Ellipsis and Gapping.

In order to develop target-like knowledge of the grammaticality contrast between VPE and Gapping in English, L1 children need to know that whereas VPE is grammatical both in conjunct clauses (e.g., (47a)) and in adjunct clauses (e.g., (47b)), Gapping is grammatical only in conjunct clauses (e.g., (47c)) and not in adjunct clauses (e.g., (47d)).

(47) a. VPE in a conjunct clause: Sara made pizza [and Kelly did [e] too].
b. VPE in an adjunct clause: Sara made pizza [because Kelly did [e]].
c. Gapping in a conjunct clause: Sara made pizza [and Kelly [e] pasta].
d. Gapping in an adjunct clause: *Sara made pizza [because Kelly [e] pasta].

(paradigm adapted from Schwartz, 1999, p. 638, (6a)–(6d))

The main challenge for L1 children is to acquire implicit knowledge that Gapping is ungrammatical in adjunct clauses. First of all, there is no positive evidence that would prevent an L1 child from allowing Gapping in adjunct clauses. Furthermore, knowledge of this ungrammaticality is not derivable from domain-general learning principles, such as analogy between VPE and Gapping. As I will show in the corpus-based study (see Chapter 4), the native input data given to L1 children contain very few instances of VPE in adjunct clauses or of Gapping. Although it is reasonable to assume that L1 children apply analogical reasoning during the language acquisition process, the case at issue—where the available input data to children are in fact quite limited—would mislead them to conclude that an illicit structure is grammatical (see Schwartz & Sprouse, 2013). Suppose that an L1 child has encountered sentences with VPE in conjunct clauses, such as (47a), and generalized this analysis/pattern to include sentences with VPE in adjunct clauses, such as (47b), as schematized in Table 2.3.

Table 2.3
*Scenario 1: Analogy from Gapping in Conjunct Clauses to *Gapping in Adjunct Clauses*

Construction	**Conjunct**		**Adjunct**
VPE	grammatical	⇢	grammatical
Gapping	grammatical	⇢	***grammatical*** **i.e., non-target-like result**

If this were what L1 children do, then one would expect them to use the same analogy to generalize from Gapping in conjunct clauses (e.g., (47c)) to Gapping in adjunct clauses

(e.g., (47d)), which would result in a non-target-like grammar.[10] We can imagine another case where an L1 child generalizes VPE in conjunct clauses to Gapping in conjunct clauses and uses the same analogy to extend VPE in adjunct clauses to Gapping in adjunct clauses, as illustrated in Table 2.4. However, this would also give rise to a non-target-like grammar. Thus, analogical reasoning cannot explain how L1 children acquire implicit knowledge of the ungrammaticality of Gapping in adjunct clauses.

Table 2.4
*Scenario 2: Analogy from VPE in Adjunct Clauses to *Gapping in Adjunct Clauses*

Construction	**Conjunct**	**Adjunct**
VPE	grammatical	grammatical
	⇣	⇣
Gapping	grammatical	***grammatical*** **i.e., non-target-like result**

Lastly, it is reasonable to assume that there is no (explicit) negative evidence provided to L1 children that would cause them to (first allow and then) disallow Gapping in adjunct clauses; in other words, L1 children do not produce sentences like (47d) and then get corrected.

There is good reason to believe that L1-Korean L2ers of English face similar learnability problems as L1-English children do when acquiring the grammaticality contrast between VPE and Gapping in English. First, L1-Korean transfer alone could not account for L1-Korean L2ers of English acquiring the implicit knowledge that VPE is grammatical in adjunct clauses. As discussed in §2.1, there is no true equivalent of VPE in Korean (e.g., H.-D. Ahn, 2018; Goldberg, 2005; J.-S. Kim, 2006; M.-K. Park, 1997). The three closest analogues of VPE in

[10] William O'Grady (personal communication, 17 April 2020) raised the possibility that learners might be more conservative in their acquisition of Gapping than in their acquisition of VPE because, unlike VPE, which often occurs in isolation (e.g., *Tom did*), Gapping does not occur in stand-alone contexts (e.g., **Tom a book*; see Chapter 4). Thus, learners might not use analogy to extend Gapping from conjunct clauses to adjunct clauses, even though they might use analogy to extend VPE from conjunct clauses to adjunct clauses. However, I do not think there is any reason to believe that learners "decide" whether or not to employ analogy based solely on the presence or absence of the construction in stand-alone contexts. Furthermore, according to Culicover & Jackendoff (2005), Gapping can appear in stand-alone clauses, as in (ix).

(ix) A: Does Robin speak French?
B: No, Lesilie, German.

(Culicover & Jackendoff, 2005, p. 276, (71d))

Korean, i.e., Argument Ellipsis (AE), *Kulay* 'Do So' Anaphora, and Pseudo-VPE, cannot occur in adjunct clauses; the two exceptions are that (a) AE is permitted to appear in adjunct clauses so long as an overt argument in the ellipsis clause has the nominative case marker (cf. *-to* 'also') and that (b) *Kulay* 'Do So' Anaphora is allowed in adjunct clauses only when its antecedent precedes it in a separate sentence or main clause. However, there is no reason to assume that L1-Korean L2ers of English initially transfer the distributional properties of Korean AE or *Kulay* 'Do So' Anaphora to English VPE.

It is also notable that the two English phenomena under discussion cannot be acquired via analogy with one another, as in the case of L1 acquisition (see above): If it were the case that L2ers generalize VPE in conjunct clauses to VPE in adjunct clauses, it should be the case that they do the same for Gapping, which would lead them to incorrectly accept Gapping in adjunct clauses. Likewise, if they generalized VPE in conjunct clauses to Gapping in conjunct clauses, they should generalize VPE in adjunct clauses to Gapping in adjunct clauses, which would also result in ungrammaticality.

Last but not least, the grammaticality contrast between VPE and Gapping cannot be learned from direct English input. To preview, the corpus-based study (see Chapter 4) found hardly any instances of VPE in adjunct clauses or of Gapping in the English as a foreign language (EFL) input to L1-Korean L2ers. VPE and Gapping are also not explicitly taught in EFL classrooms in Korea because neither of them has been targeted in the Korean National Curriculum (Ministry of Education, 2015).

2.4.2 Interpretation contrast between VP-Ellipsis and Gapping.

The paradigm for the interpretation contrast between VPE and Gapping in English differs from that for the grammaticality contrast between the two phenomena. Whereas the ungrammaticality involves Gapping in the grammaticality contrast paradigm, the ungrammaticality involves VPE in the interpretation contrast paradigm: In contrast to Gapping sentences that allow both the subject reading and the object reading, as in (48), VPE sentences permit only the SR and not the OR, as in (49).

(48) Gapping: Mom hugged the boy at home and Dad [e] in the park[a] / and [e] Dad in the park[b].

a. SR: 'Mom hugged the boy at home, and Dad (hugged the boy) in the park.'

b. OR: 'Mom hugged the boy at home, and (Mom hugged) Dad in the park.'

(49) VPE: Mom hugged the boy at home and Dad did [e] too.

a. SR: 'Mom hugged the boy at home, and Dad hugged the boy at home.'

b. OR: *'Mom hugged the boy at home, and Mom hugged Dad at home.'

Regarding the contrast at issue, the key challenge for L1-English children is to know that VPE with an OR is impossible. There is no positive evidence that could indicate to the L1 child the ungrammaticality of the OR for VPE. Analogical reasoning would also lead to the wrong result, like with the VPE–Gapping grammaticality contrast (this time, however, going from Gapping, where both readings are possible, to VPE). In addition, L1 children do not receive (explicit) negative evidence that would enable them to work out that the OR is prohibited in VPE.

The interpretation contrasts at issue constitute learnability problems also for L1-Korean L2ers. Specifically, the fact that VPE disallows the OR cannot come from these L2ers' L1. This is because (a) Korean does not have an exact analogue to English VPE and (b) AE, *Kulay* 'Do So' Anaphora, and Pseudo-VPE actually all allow the OR (see §2.1). Furthermore, the target contrasts are not derivable from direct English input; the EFL input to L1-Korean L2ers analyzed in the corpus-based study (see Chapter 4) exhibited very few instances of VPE in adjunct clauses and of Gapping. Lastly, neither of VPE nor of Gapping is a topic of classroom instruction in Korea (Ministry of Education, 2015).

2.5 Summary of Chapter

This chapter provided a syntactic analysis of VPE and Gapping in English and Korean. English has VPE, which I take to involve VP-deletion under the condition of identity with another VP in the given syntactic or discourse context (e.g., Chomsky, 1995; Fiengo & May, 1994; Hankamer & Sag, 1976). In contrast, Korean does not have VPE (e.g., H.-D. Ahn, 2018; Goldberg, 2005; J.-S. Kim, 2006; M.-K. Park, 1997). I provided evidence that AE, *Kulay* 'Do So' Anaphora, and Pseudo-VPE cannot be considered direct counterparts of English VPE. Gapping, however, is present in both English and Korean although these two languages differ in

the direction of Gapping. Adopting Johnson's (2000, 2006, 2009) analysis, I assume that this phenomenon results from ATB movement of VPs.

I also discussed the grammaticality and interpretation contrasts between VPE and Gapping in English, which are the target phenomena in this dissertation. These contrasts have bearing on the issue of learnability in L1 acquisition because of the lack of both positive evidence and (explicit) negative evidence (Crain, 1991). This is also the case for Korean speakers' L2 acquisition of English because the contrasts at issue are (a) not present in the L1 grammar, (b) not learnable from the input alone, and (c) not taught (Schwartz & Sprouse, 2000). Moreover, in the context of both L1 and L2 acquisition of English, I showed that analogical extension—from VPE to Gapping and from Gapping to VPE—is not helpful, either. This issue of learnability is the chief motivation of the current dissertation: to test for knowledge of the discussed contrasts in L1-English children and in early and late L1-Korean L2ers of English, which has not been done previously.

CHAPTER III
ACQUISITION/PROCESSING RESEARCH ON VP-ELLIPSIS AND GAPPING

This chapter looks at previous acquisition/processing research on VP-Ellipsis (VPE) and Gapping in English. Sections 3.1 and 3.2 examine L1 and L2 acquisition studies on VPE, respectively; these studies have mainly investigated pronoun interpretation and parallelism effects. Section 3.3 reviews previous studies on the L2 acquisition of the direction of Gapping. This is followed by a review of prior work on adult native processing of English Gapping. Section 3.5 summarizes the chapter. Finally, §3.6 lists the research questions for this dissertation.

3.1 First Language Acquisition Research on VP-Ellipsis

VPE has been the subject of great interest in L1 acquisition research, with a particular focus on the interpretation of pronouns (e.g., Foley et al., 2003; Matsuo, 2007; Thornton & Wexler, 1999). For example, there are different possibilities for the interpretation relation of the overt pronoun in the antecedent clause and the elided one in the ellipsis clause in (1).

(1) Oscar bites his apple and Bert does [e] too.

As noted in §2.1, a sloppy reading and strict readings are both available in this type of sentence, as shown in (2a)–(2d).

(2) Interpretations

Sloppy:

a. O bites O's apple and B bites B's apple. ii jj

Strict:

b. O bites O's apple and B bites O's apple. ii ji

c. O bites B's apple and B bites B's apple. ij jj

d. O bites E's apple and B bites E's apple. ik jk

(Foley et al., 2003, p. 53, (1))

However, there are a variety of other logical interpretation relations of the overt pronoun and the elided one in (1) that are prohibited under both sloppy and strict readings, as shown in (3).

(3) Ungrammatical interpretations:

a. *O bites O's apple and B bites E's apple.	ii	jk
b. *O bites B's apple and B bites O's apple.	ij	ji
c. *O bites B's apple and B bites E's apple.	ij	jk
d. *O bites E's apple and B bites O's apple.	ik	ji
e. *O bites E's apple and B bites B's apple.	ik	jj

(Foley et al., 2003, p. 53, (1'))

Such constrained ambiguity associated with pronouns in VPE constitutes a learnability problem for L1 acquisition because there is no positive evidence that prevents children from ruling out the impossible interpretations. Despite this challenge, L1 acquisition research has shown that children accept the possible interpretations and reject the impossible ones (e.g., Foley et al., 2003; Matsuo, 2007).

For example, Foley et al. (2003) investigated the interpretation of possessive pronouns inside the ellipsis clause using sentences like (1). Children aged 3;0−7;11 completed an act-out task in which they were asked to use a set of toys to show what a target sentence means, which was preceded by the story. They also performed a picture-sentence matching task where they matched a sentence to the corresponding picture. Despite the learnability challenges discussed above, the children succeeded in accepting only the subset of possible interpretations of the pronouns (e.g., (2)), although they preferred the sloppy interpretation (e.g., (2a)). Importantly, the children never acted out or selected ungrammatical interpretations (e.g., (3)).

Matsuo (2007) obtained similar findings with a truth-value judgment task (TVJT). The stimuli for her TVJT comprised two possible conditions (sloppy and strict) and two impossible conditions ("color mismatch" and "object mismatch"), as exemplified in (4a)–(4d). For example, (4c) shows that the target sentence, *Mr. Bear found a blue fish and Mr. Tiger did too*, has an ungrammatical interpretation given the scenario in which Mr. Bear found a blue fish and Mr. Tiger found a pink fish (i.e., color mismatch). The target sentence in (4d), which shows an object mismatch, is also ungrammatical on the intended interpretation. L1-English children aged

4;9−6;9 and L1-English adults both accepted possible interpretations and rejected impossible ones.

(4) a. Condition 1: Sloppy reading

Scene: Cookie Monster ate Cookie Monster's cookie and Mike ate Mike's cookie.

Target: Cookie Monster ate his cookie and Mike did too.

b. Condition 2: Strict reading

Scene: The mother hid behind the mother's tree and the girl hid behind the mother's tree.

Target: The mother hid behind her tree and the girl did too.

c. Condition 3: Color mismatch

Scene: Mr. Bear found a blue fish and Mr. Tiger found a pink fish.

Target: *Mr. Bear found a blue fish and Mr. Tiger did too.

d. Condition 4: Object mismatch

Scene: The cow ate asparagus and the elephant ate carrots.

Target: *The cow ate asparagus and the elephant did too.

(adapted from Matsuo, 2007, pp. 10–13, (12)–(15))

As such, children's early knowledge of the possible and impossible ambiguity of pronouns unrealized in the ellipsis clause has been tested and confirmed in previous research. Other evidence pointing toward children's full mastery of VPE comes from studies that examined other VPE effects, such as parallelism effects (e.g., Matsuo & Duffield, 2001; Murphy, 1985a, 1985b). The focus of these studies was on comparing VPE (e.g., (5a)) and VP-Anaphora (VPA; e.g., (5b)); in the latter, a form of *do x* (e.g., *do it/that/so*) replaces a vP in an anaphoric relation with its corresponding constituent in the antecedent clause (Houser, 2010; Stroik, 2001). Specifically, VPE exhibits a firm structural parallelism effect. If VPE does not have a syntactically identical antecedent, as in the case in passive-active pairs (e.g., (5c) vs. (5a)) and nominal-verbal pairs (e.g., (6c) vs. (6a)), it results in ungrammaticality, or at least (highly) degraded acceptability. Such syntactic parallelism effects are not observed in VPA, as shown in (5b) vs. (5d) and in (6b) vs. (6d).

(5) a. Someone had to fix the cat's tire. Cookie Monster said he was able to. (Active-VPE)

b. Someone had to fix the cat's tire. Cookie Monster said he was able to do it. (Active-VPA)

c. *The cat's tire had to be fixed. Cookie Monster said he was able to. (Passive-VPE)

d. The cat's tire had to be fixed. Cookie Monster said he was able to do it. (Passive-VPA)

(adapted from Matsuo & Duffield, 2001, p. 311, Table 1, (2))

(6) a. The clowns were very full, and the dogs said that they should take a nap.
But the clowns didn't want to. (VP-VPE)

b. The clowns were very full, and the dogs said that they should take a nap.
But the clowns didn't want to do that. (VP-VPA)

c. *The clowns were very full, and the dogs said that a nap would be good for them.
But the clowns didn't want to. (NP-VPE)

d. The clowns were very full, and the dogs said that a nap would be good for them.
But the clowns didn't want to do that. (NP-VPA)

(adapted from Matsuo & Duffield, 2001, p. 311, Table 1, (9))

On the other hand, when involved in antecedent contained deletion (ACD), VPA is ungrammatical but VPE is grammatical, as shown in (7b) vs. (7a). Such a grammaticality contrast is explained using the bound variable constraint (Fiengo & May, 1994; Wasow, 1972), which bars a relative clause operator before *that* in the CP from binding the variable *so*. Note that if VPE and VPA are conjoined as in (7c) and (7d), both are grammatical.

(7) a. The girls baked the same cookies that the cats did. (VPE-ACD)

b. *The girls baked the same cookies that the cats did so. (VPA-ACD)

c. The girls baked some cookies and the cats did too. (VPE-conjoined)

d. The girls baked some cookies and the cats did so too. (VPA-conjoined)

(Matsuo & Duffield, 2001, p. 326, Appendix, (8))

These contrasts between VPE and VPA involve a learnability challenge. Specifically, the absence of (ungrammatical) sentence types, such as (5c), (6c), and (7b), in the input cannot be

taken by children as evidence of their ungrammaticality. Furthermore, the parallelism constraint applies only to VPE and the bound variable constraint applies only to VPA. Given that VPE and VPA are, it seems, used in semantically (and pragmatically) equivalent contexts, it is reasonable to think that children would have difficulty making the target distinctions between VPE and VPA, that is, if the sole source of these distinctions were input.

To test for knowledge of VPE and VPA, Matsuo and Duffield (2001) administered a contextualized acceptability judgment task to children aged 3;11−6;7. All children received four sentences testing the active vs. passive contrast (e.g., (5)), four testing the VP vs. NP contrast (e.g., (6)), and four testing the ACD vs. conjoined contrast (e.g., (7)). The results showed that the children were able to accurately judge the grammaticality of the experimental sentences, respecting both the structural parallelism constraint on VPE and the bound variable constraint on VPA. These results clearly indicate the presence of knowledge of VPE and VPA in children as young as 3;11.

In sum, despite the learnability problems, young (English-acquiring) children successfully precluded the impossible interpretations of elided pronouns in VPE as well as violations of the constraints on VPE and VPA. As for the L2 acquisition of English, the learnability challenge again holds in the case where the L2ers' native language behaves differently from English with respect to the phenomena at issue. Only a few L2 studies have examined this, and I turn to them in the following section.

3.2 Second Language Acquisition Research on VP-Ellipsis

Duffield and Matsuo (2009) explored sensitivity to parallelism effects in the L2 acquisition of VPE vs. VPA by testing L2ers from diverse L1 backgrounds (Dutch: $n = 20$, Japanese: $n = 19$, Spanish: $n = 20$). The differences between the L2ers' native languages and English regarding VPE and VPA are summarized in Table 3.1.[1]

[1] Duffield and Matsuo (2009, p. 119, fn. 4) note that the acceptability judgments of the language data in Dutch, Japanese, and Spanish were obtained from previous linguistic research (e.g., Carretero, 1999; Lobeck, 1995; Lopez, 1994; Otani & Whitman, 1991; Santos, 2006) or from linguists who are native speakers of the relevant languages.

Table 3.1
Crosslinguistic Differences in VPE, VPA, NOC, and NCA

	Antecedent sentence type			
Construction type	**Active**	**Passive**	**Verbal**	**Nominal**
English				
VPE	✓	*	✓	*
VPA	✓	✓	✓	✓
Dutch				
VPE	NA	NA	NA	NA
VPA	✓	??	✓	*
Japanese				
VPE	NA	NA	NA	NA
NOC	✓	✓	✓	*
VPA	✓	?	✓	?
Spanish				
VPE	NA	NA	NA	NA
NCA[2]	✓	*	✓	*
VPA	✓	✓	✓	*

Notes. Adapted from Duffield & Matsuo, 2009, p. 104, Table 1. NCA = Null Complement Anaphora; NOC = Null Object Construction.

Dutch does not have a direct counterpart of English VPE, but it does have VPA, which exhibits a stronger parallelism effect for nominal antecedents than for passive ones.[3] According to Duffield and Matsuo (2009), the closest construction to English VPE in Japanese is the NOC,

[2] Null Complement Anaphora (NCA) involves TP-ellipsis licensed by modal predicates (Dagnac, 2010), as shown in (i).

(i) Tom pudo ver a Lee, pero María no pudo [e].
Tom could see to Lee, but Mary NEG could
'Tom could see Lee but Mary couldn't.'

(Dagnac, 2010, p. 158, (3b))

[3] (iia) and (iib), respectively, show that Dutch VPA is degraded with nonparallel passive antecedents (Duffield & Matsuo, 2009) and unacceptable with nonparallel nominal antecedents (personal communication with a Dutch speaker, 4 February 2019).

(ii) a. ? Het vuilnis moest buiten gezet worden, maar Marie wilde het niet doen.
The garbage had to out set be but Marie wanted it NEG do
'The garbage had to be taken out, but Marie didn't want to do it.'

(Duffield & Matsuo, 2009, p. 102, (9))

b. * John wilde een kus, maar Marie wilde het niet doen.
John wanted a kiss but Marie wanted it NEG do
'John wanted a kiss, but Mary didn't want to do it.'

in which nonparallel passive antecedents are acceptable but nominal antecedents are unacceptable.[4] Japanese exhibits VPA with mild parallelism effects.[5] In Spanish, the closest phenomenon to VPE is Null Complement Anaphora (NCA), which complies with the same parallelism constraint as in English VPE.[6] Spanish also manifests a form of VPA that is compatible with passive antecedents but incompatible with nominal antecedents.[7]

---

[4] Japanese NOCs are compatible with nonparallel passive antecedents, as in (iiia), but incompatible with nonparallel nominal antecedents, as in (iiib).

(iii) a. Gomi-wa das-sare-nakerebanaranakat-ta, demo watasi-wa dasi-taku-nakat-ta.
garbage-TOP take out-PASS-have to-PST but I-TOP take out-want-NEG-PST
'??The garbage had to be taken out, but I didn't want to (take it out).'

b. * John-wa kisu-ga hosikat-ta ga, Mary-wa si-taku-nakat-ta.
John-TOP kiss-NOM want-PST but Mary-TOP do-want-NEG-PST
'*John wanted a kiss but Mary didn't want to.'

(adapted from Duffield & Matsuo, 2009, p. 102, (11))

[5] Japanese VPA is degraded with nonparallel passive and nominal antecedents, as in (iva) and (ivb).

(iv) a. ? Gomi-wa das-sare-nakerebanaranakat-ta, demo watasi-wa soo si-taku-nakat-ta.
garbage-TOP take out-PASS-have to-PST but I-TOP so do-want-NEG-PST
'The garbage had to be taken out, but I didn't want to do so.'

b. ? John-wa kisu-ga hosikat-ta ga, Mary-wa soo si-taku-nakat-ta.
John-TOP kiss-NOM want-PST but Mary-TOP so do-want-NEG-PST
'John wanted a kiss but Mary didn't want to do so.'

(adapted from Duffield & Matsuo, 2009, p. 102, (12))

[6] Spanish NCA is subject to parallelism constraints, like English VPE is. For example, it is ungrammatical when it has a nonparallel passive antecedent, as in (va), or a nonparallel nominal antecedent, as in (vb).

(v) a. * La basura tiene que sacarse, pero yo no quiero.
The garbage has COMP take out but I NEG want
'*The garbage has to be taken out, but I don't want to.'

b. * Juan quería un beso, pero Maria no quería.
John wanted a kiss but Mary NEG wanted
'*John wanted a kiss, but Mary didn't want (to).'

(Duffield & Matsuo, 2009, p. 103, (13))

[7] Compare (via) and (vib) for the grammaticality contrast between Spanish VPA with a nonparallel passive antecedent and that with a nonparallel nominal antecedent.

(vi) a. La basura tiene que sacarse, pero yo no quiero hacerlo.
the garbage has COMP take out-REFL but I NEG want do-it
'The garbage has to be taken out, but I don't want to do it.'

b. * Juan quería un beso, pero Maria no quería hacerlo.
John wanted a kiss but Mary NEG wanted do-it
'John wanted a kiss, but Mary didn't want to do it.'

(Duffield & Matsuo, 2009, p. 103, (14))

Duffield and Matsuo (2009) point out that the effect of parallelism on VPE and VPA in English is "a subtle, language-particular, and construction-specific property, making its acquisition by L2 learners especially challenging" (p. 104). That is, based solely on the input, it would be difficult/impossible for L2ers to acquire implicit knowledge that VPE is sensitive to the syntactic structure of the antecedent clause but VPA is not. In addition, none of the L2ers' native languages (Dutch, Japanese, Spanish) in Duffield and Matsuo's study works similarly to English concerning the structural parallelism effects in VPE vs. VPA.

Duffield and Matsuo (2009) created a sentence completion judgment task in which the L2ers judged "whether the target sentence [was] a sensible and accurate completion of the [context sentence]" (p. 312) by pressing the *yes* or *no* button on a button box connected to the computer. The participants' L2 English proficiency was not independently assessed, nor was their background information (e.g., length of TL exposure) provided. The results revealed reliable differences among the various L1 groups. The L1-Dutch L2ers, like the English native speakers, displayed significantly stronger parallelism effects for VPE than for VPA, despite the absence of VPE in their L1. This result was interpreted by Duffield and Matsuo as indicating that the L1-Dutch L2ers had converged on the relevant aspects of English grammar. However, Duffield and Matsuo also observed that many of the L1-Japanese and L1-Spanish L2ers showed signs of L1 transfer (p. 115). These findings might suggest that full attainment of the target phenomena depends on the learner's L1.

Caution is required, however, in drawing conclusions about the possibility of convergence on the TL grammar depending on properties of the L1. Even though the L1-Dutch L2ers accepted (a) VPE with active antecedents (e.g., (5a)) significantly more often than VPE with passive antecedents (e.g., (5c)) and (b) VPE with verbal antecedents (e.g., (6a)) significantly more often than VPE with nominal antecedents (e.g., (6c)), their acceptance rates for the passive antecedent type (i.e., 74%) and the nominal antecedent type (i.e., 68%) were relatively high, thereby indicating that the L2ers, as a group, did not fully reject these nonparallel sentence types. Even if it were the case, as concluded by Duffield and Matsuo (2009, p. 115), that the performance of the L1-Dutch mirrored that of the English native speakers, it is unclear what led only this group (vs. the L1-Japanese and L1-Spanish groups) to target-like knowledge of VPE and VPA given that all three L1s operate differently from English (see Table 3.1). Furthermore, because Duffield and Matsuo did not test for L2 proficiency or report important

background information about the L2ers (e.g., length of TL exposure), it is not known whether these factors might have contributed to the results.

L1 effects for VPE in L2 acquisition were later investigated by Hawkins (2012). He claimed that the grammaticality distribution of VPE in English (e.g., (8) vs. (9)) constituted a learnability problem for his L2 participants, whose L1s were Arabic (*n* = 19) or Mandarin (*n* = 20). Neither Arabic nor Mandarin has the same type of VPE as English.[8] Moreover, input alone cannot lead L2ers to know that the sentences in (8) are grammatical/felicitous but those in (9) are ungrammatical/infelicitous. According to Hawkins, learners may encounter sentences showing the pattern *lexical verb* (e.g., *wrote*) … *auxiliary/modal verb* (e.g., *did*) + *ellipsis*, as in (8a)−(8c). However, there is nothing in the input to let them know that the other cases displaying the same pattern, such as (9b), are infelicitous.

(8) a. Jack wrote Jill a letter. Mary did [e] too.
b. Jack wrote Jill a letter. Mary has [e] too.
c. Jack wrote Jill a letter. Mary will [e] too.

(adapted from Hawkins, 2012, p. 405, (1))

[8] According to Hawkins (2012, p. 416), VPE is not licensed at all in Arabic. Chinese also does not have English-like VPE, but presents two similar constructions. One construction involves ellipsis after copula *shi* 'be,' as in (vii).

(vii) Zhangsan kanjian-le tade mama, Lisi ye shi [e].
Zhangsan see-ASP his mother, Lisi also is
'Zhangsan saw his mother; Lisi did too.'

(Hawkins, 2012, p. 417, (29))

The other is the Null Object Construction where a lexical verb in the antecedent clause is repeated in the ellipsis clause, as shown in (viii).

(viii) Zhangsan kanjian-le tade mama, Lisi ye kanjian-le [e].
Zhangsan see-ASP his mother, Lisi also see-ASP
'*Zhangsan saw his mother; Lisi saw too.'

(Hawkins, 2012, p. 417, (30))

(9) a. Jack sent Jill a letter. *Mary sent [e] too.
b. Jack wrote Jill a letter. #Mary was [e] too.
c. Jill is very successful. #Mary will too.

(Hawkins, 2012, p. 405, (2))

In addition, positive evidence available in the input does not tell L2ers that VPE is ungrammatical or degraded if its antecedent is not syntactically parallel, as shown in (10b) vs. (10a).

(10) a. None of the astronomers saw the comet, but John did.
b. ??The comet was nearly unseeable, but John did.

(adapted from Hawkins, 2012, p. 420, Table 1, (10a) & (10c))

Hawkins (2012) used the same sentence completion judgment task as Duffield and Matsuo (2009) did, but the task had different response options: "perfect," "possible," and "impossible." Importantly, he also added an independent proficiency test. He found that all intermediate and advanced L1-Arabic L2ers of English and L1-Mandarin L2ers of English were able to distinguish parallel from nonparallel VPE sentences ($k = 16$; e.g., (10)), grammatical from ungrammatical VPE sentences ($k = 6$; e.g., (11)), and felicitous from infelicitous VPE sentences ($k = 21$; e.g., (12)).

(11) a. It was Jill's birthday. John sent her a card by email. Tom thought that Mary did too.
b. Jill and Mary were applying for the same job. Jill sent an application by email. *John thought that Mary sent too.

(adapted from Hawkins, 2012, p. 420, Table 1, (2)–(3))

(12) a. Sue has sold her house. John believed that Mary did too.
b. #Jill is very successful. Tom thinks that Sue can too.

(adapted from Hawkins, 2012, p. 420, Table 1, (5), (9))

Except for the type of sentence shown in (8b), the L2ers in Hawkins's (2012) study were able to make the same syntactic distinctions as native speakers did in their treatment of English VPE (for native speakers' judgments, see Arregui et al., 2006; Tanenhaus & Carlson, 1990), regardless of L1 background (Arabic or Mandarin) or regardless of proficiency level (intermediate or advanced). Hawkins claimed that such (near) native-like performance provides evidence for L2 knowledge that does not come from the L1 grammar or from positive evidence in the TL input.

Hawkins (2012) states that his L2ers' failure to accept (8b) is consistent with the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007). This hypothesis posits that L2ers have difficulty acquiring uninterpretable features that are not instantiated in their L1, such as, according to Rouveret (2010), the perfective feature on *-en* elided in (8b). Adopting this analysis, Hawkins assumes that *-en* involves an uninterpretable feature because it can occur either in perfective verb forms (e.g., *have eaten*) or in passive forms (e.g., *be eaten*). However, ellipsis of this uninterpretable feature in (8b) does not result in ungrammaticality because the perfective meaning can be easily recovered from the stranded auxiliary *have*. This explains the grammaticality of both (13a) and (13b), which in fact differ with respect to the identity of the verb form between the antecedent and ellipsis clauses; even when its antecedent clause is in the simple present tense, hence non-identical to the ellipsis clause, (13b) is grammatical.

(13) a. Susan has retired from teaching, and Wendy has ~~retired from teaching~~ too.
(Perfective auxiliary stranding; Identical antecedent verb)

b. Tom writes poetry, and Bill has ~~written poetry~~ too.
(Perfective auxiliary stranding; Non-identical antecedent verb)

c. Caroline is running fast, and Heather is ~~running fast~~ too.
(Progressive auxiliary stranding; Identical antecedent verb)

d. *Julie sleeps late, and Debbie is ~~sleeping late~~ too.
(Progressive auxiliary stranding; Non-identical antecedent verb)
(Al-Thubaiti, 2019, p. 251, Appendix I, (35), (39), (27), (31))

In contrast to the perfective feature on *-en*, the progressive feature on *-ing* is interpretable. When it undergoes deletion from the ellipsis site after the non-identical antecedent, as in (13d),

this results in ungrammaticality. In this case, there is no way to recover the progressive meaning from the stranded auxiliary *be*, which can encode different meanings in various constructions, such as attributive meaning (e.g., *be happy*), progressive meaning (e.g., *be eating*), and stative meaning in passive constructions (e.g., *be eaten*). Note that (13c), where the antecedent and the elided verbs are morphologically identical, is grammatical.

Al-Thubaiti (2019) further investigated the subtle contrasts associated with VPE in a study involving 15 native English controls and 34 L1-Saudi Arabic L2ers of English. The L2ers were divided into two proficiency groups based on the results of an independent proficiency test: advanced (n = 19) and very advanced (n = 15). They completed an acceptability judgment task using a 5-point scale ranging from 1 (definitely impossible) to 5 (definitely possible). There were 164 stimuli in total, including 48 critical sentences, which were presented both aurally and visually.

The critical items comprised three sets of VPE sentences. One set compared perfective *have* stranding and progressive *be* stranding with identical and non-identical antecedents, as shown in (13) above. The second set consisted of modal-stranding constructions that involved ellipsis of the non-finite copula *be* (e.g. (14a), (14b)) and lexical verbs (e.g., (14c)–(14f)); each came with either an identical antecedent of the elided verb (e.g., (14a), (14c)) or a non-identical antecedent of the elided verb (e.g., (14b), (14d), (14e), (14f)).

(14) a. John will be here, and Mary will ~~be here~~ too.

(Copula; Identical antecedent verb)

b. * John is happy, and Mary will ~~be happy~~ soon.

(Copula; Non-identical antecedent verb)

c. Bruce can design a webpage, and George can ~~design a webpage~~ too.

(Lexical verb; Identical antecedent verb)

d. Mike revises his work, and Jennifer should ~~revise her work~~ as well.

(Lexical verb; Non-identical antecedent verb: Simple finite)

e. ? Bill is writing an essay, and Tom should ~~write an essay~~ as well.[9]

(Lexical verb; Non-identical antecedent verb: Progressive participle)

f. John has baked a cake, and Mary will ~~bake a cake~~ too.

(Lexical verb; Non-identical antecedent verb: Perfective participle)

(adapted from Al-Thubaiti, 2019, p. 234, Table 2, (1))

According to Lasnik (1999; see also Rouveret, 2010, 2012), the contrast in acceptability between (14b) and (14d) stems from when in the derivation of a sentence verbal inflection occurs; while lexical verbs are inflected at the syntax level through morphological merge, the copula is inflected in the lexicon before entering the syntactic derivation. This allows lexical verbs, but not the copula, to ignore morphological parallelism. The first and second sets had four items per sentence type. Each sentence type was counter-balanced in terms of the presence of negation after the stranded auxiliaries or modals. For example, among the four experimental items for the type in (13b) (viz. "Perfective auxiliary stranding; Non-identical antecedent verb" type), two did not have negation after the auxiliary, as in (13b), and the other two did, as illustrated in (15).

(15) John plays football, but Tom hasn't ~~played football~~ recently.

(Perfective auxiliary stranding; Non-identical antecedent verb)

(Al-Thubaiti, 2019, p. 234, Table 2, (2))

[9] Lasnik (1999) considers (14e) to be only marginally acceptable. The reason for this degraded acceptability is unknown.

The last set of critical stimuli included eight ungrammatical items that have a stranded lexical verb without its object, as shown in (16).

(16) *Neal grows vegetables, and his wife grows ~~vegetables~~ too.

(Al-Thubaiti, 2019, p. 234, Table 2, (3))

Al-Thubaiti (2019) argued that the above grammaticality contrasts constituted a poverty-of-the-stimulus problem for her L1-Saudi Arabic L2ers because they are underdetermined by the input and not taught in language classes (p. 226). Furthermore, Saudi Arabic does not license VPE (Abdulkarim & Roeper, 1997).

In the acceptability judgment task, both L2 proficiency groups showed target-like performance on the contrast in (13), (14), and (16) in general. They were able to accept (13a) and (13c) and reject (13d), just as the native controls did. However, they failed to accept (13b) with ellipsis of the uninterpretable feature on *-en*, which is in line with the results in Hawkins (2012). Al-Thubaiti's (2019) follow-up *t*-test analyses on the sentence type in (13b) revealed that while the native speakers' ratings were significantly higher than the mid-point rating, the very advanced group's ratings were not significantly different from the mid-point rating, and the advanced group's ratings were significantly lower than the mid-point rating.

In her further analysis of this sentence type, Al-Thubaiti (2019) did find a significant effect of the presence of negation in the ellipsis clause in both proficiency groups; their ratings were significantly higher for perfective *have* stranding with negation (e.g., (15)) than without negation (e.g., (13b)). In addition, the very advanced group's ratings for perfective *have* stranding with negation ($M = 3.67$; $SD = 0.90$) were comparable to the English native speakers' ratings ($M = 3.93$; $SD = 0.70$). On the basis of these results, Al-Thubaiti concluded that very advanced L2ers can overcome the poverty-of-the-stimulus problem, evincing the subtle contrast between *-ing* and *-en* at the elliptical site, but only by resorting to another interpretable feature, namely, negation (p. 246)—a conclusion which is again compatible with the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007).

This conclusion has to be taken with caution, however. In light of Al-Thubaiti's (2019) admission that there was "much individual variation" (p. 240) in the very advanced group, it is conceivable that at least some L2ers in the group exhibited a native-like performance pattern on

perfective *have* stranding both with and without negation. This group's mean acceptability rating of perfective *have* stranding without negation was 2.77 (out of 5), with a large standard deviation of 1.59 (see Al-Thubaiti, 2019, p. 242, Table 7).

The only existing work on L1-Korean L2ers' knowledge of VPE in English is a study conducted by J. Kim (2012, 2015). Her 30 participants, who had diverse levels of English proficiency, all demonstrated the ability to comprehend VPE in a target-like manner even though there is no direct equivalent in their L1. To explore how VPE is interpreted, three different story contexts were created for an offline questionnaire following Matsuo (2007): (a) a "full-match" context where two out of three main characters performed the action with exactly the same objects ($k = 4$); (b) a "color-mismatch" context in which two of the three characters did the same action using the same type of objects but with different colors ($k = 4$); (c) an "object-mismatch" context where two out of the three characters performed the same action, using totally different kinds of objects ($k = 4$). Participants were asked to read short stories and then judge whether the subsequent sentence matched the story or not. The results showed that in the full-match context, L1-Korean L2ers approached near-ceiling performance, accepting VPE 97.5% of time. By contrast, no L2ers accepted VPE in the two mismatch contexts. This finding that L2ers assigned appropriate interpretations to VPE sentences is in line with Matsuo's results for L1-English children.

In sum, the findings from previous studies on VPE in L2 English have not been entirely consistent. First, the L2ers in the studies by Al-Thubaiti (2019), Hawkins (2012), and J. Kim (2012, 2015) generally achieved native-like performance except on the sentence type with stranded auxiliary *have*, e.g., in (8b) and (13b). Second, the L1-Japanese and L1-Spanish L2ers (as well as, arguably, the L1-Dutch L2ers) in Duffield and Matsuo's (2009) study did not perform like native speakers. Because there may have been individual L2ers in the Hawkins and Al-Thubaiti studies who did have target-like performance on all VPE sentence types and because we do not know whether the results in Duffield and Matsuo's study were attributable to unreported important factors, such as L2 proficiency or length of TL exposure, it is hard to draw firm conclusions at this point. Furthermore, empirical findings related to VPE have so far been limited to elided pronouns and parallelism effects (relative to VPA). This dissertation, by contrast, will focus on the contrast between VPE and Gapping. The next section discusses empirical studies on the L2 acquisition of Gapping.

3.3 Second Language Acquisition Research on Gapping

To the best of my knowledge, there have been no L1 acquisition studies on Gapping and only very few in L2 acquisition research (e.g., Kanno, 1999; O'Grady, 1999). L2ers' knowledge of Gapping in English and Japanese was first examined by O'Grady (1999), who proposed that a language's head-complement order constrains a particular Gapping direction, as laid out in (17).

(17) Constraint on Gapping Direction

a. Verb-Object languages (e.g., English): *[S … Ø …] [S … V …]
b. Object-Verb languages (e.g., Japanese): *[S …… V] [S …… Ø]

(adapted from O'Grady, 1999, p. 143, (6))

On the basis of the constraint in (17) and his intuition that Gapping is hardly ever provided as input, O'Grady made the following predictions:

> If learners succeed in using word order to reject the inappropriate gapping pattern, we can attribute their success to the operation of the acquisition device itself rather than to instruction or direct experience. On the other, if they fail, we can draw inferences from this fact about possible deficits in the acquisition device that is available for post-adolescent second language acquisition. (p. 144)

To test these predictions, four groups of participants were recruited: (a) L1-Japanese L2ers of English (n = 34); (b) L1-English L2ers of Japanese (n = 75); (c) native speakers of English (n = 10); and (d) native speakers of Japanese (n = 10). All four groups completed a written acceptability judgment task using a 5-point scale. There was one condition with forward Gapping and another with backward Gapping (k = 5 per condition). The experimental sentences for English and Japanese are illustrated in (18) and (19), respectively.

(18) a. Forward Gapping:

[John reads Time] and [Sue [e] Newsweek].

b. Backward Gapping:

* [John [e] Time] and [Sue reads Newsweek].

(O'Grady, 1999, p. 142, (1))

(19) a. Forward Gapping:

* [John-wa Time-o yon-de] [Sue-wa Newsweek-o [e]].
[John-TOP Time-ACC read-GER] [Sue-TOP Newsweek-ACC]
'John read Time and Sue Newsweek.'

b. Backward Gapping:

[John-wa Time-o [e]] [Sue-wa Newsweek-o yon-da].
[John-TOP Time-ACC] [Sue-TOP Newsweek-ACC read-PST]
'John read Time and Sue Newsweek.'

(O'Grady, 1999, p. 142, (2))

The results showed that while the L1-Japanese L2ers of English accurately rejected backward Gapping in English, they rated the grammatical forward Gapping pattern around the middle point of the scale. However, the L1-English L2ers of Japanese did not judge the Japanese sentences in a target-like manner. Instead, they favored the (ungrammatical) forward Gapping pattern.

A similar experiment was conducted by Kanno (1999) to test the Constraint on Gapping Direction in L2 Japanese. This time, in addition to L1-English L2ers of Japanese (n = 92), there were also L1-Mandarin L2ers of Japanese (n = 75) whose L1 does not have Gapping at all (O'Grady, 1999). The experimental design was the same as that in O'Grady except that all sentences were preceded by a context passage. Consistent with O'Grady's findings, the L1-English group failed to reject the ungrammatical forward Gapping pattern in Japanese, which was actually rated higher than the grammatical backward pattern. The L1-Mandarin group successfully rejected the ungrammatical forward Gapping pattern in Japanese, but this group did not show a clear tendency for the grammatical backward Gapping pattern.

Kanno (1999) explains these results as an L1 effect. She raises the possibility that the absence of Gapping in Mandarin allowed the L1-Mandarin L2ers of Japanese to be free from any

"transfer-based prejudices" (p. 171) that can be attributed to their L1 grammar. This might have led this group's acquisition device to operate in the same way as it would in L1 acquisition. In contrast, the L1-English L2ers of Japanese began at a disadvantage, since the Gapping direction is opposite in their L1.

However, it should be pointed out that L1 transfer arguably works the same regardless of whether an L1 does or does not manifest the same phenomenon as in the TL (e.g., Schwartz, 1997).[10] From this perspective, it is more reasonable to assume that both groups were under the influence of "transfer-based prejudices," but the difference in their L1s led to the different results. Because there is no Gapping in Mandarin, the L1-Mandarin L2ers came to the task with no preferred Gapping direction, which may have made it easier for them to acquire the pattern of Japanese. On the other hand, L1-English L2ers would have come to the task allowing forward Gapping direction and prohibiting backward Gapping direction, which hindered their acquisition of the Gapping pattern in Japanese.

In sum, the L2 data in both O'Grady (1999) and Kanno (1999) showed a lack of target-like performance (by groups). Although O'Grady's L1-Japanese L2ers of English and Kanno's L1-Mandarin L2ers of Japanese were able to reject the ungrammatical Gapping pattern, they failed to accept the grammatical Gapping pattern in the TL. This might be due to short exposure (Kanno, 1999, p. 167; O'Grady, 1999, p. 152) or low proficiency. Given that neither study reported L2 participant information about either, it remains speculative whether one or both factors contributed to the results. This dissertation attempts to partially address this issue by testing for proficiency effects among L2ers by means of an independent proficiency measure.

In addition, previous L2 research on VPE and Gapping has focused solely on adults. This dissertation thus aims to expand the current body of research by testing both child ("early") and adult ("late") L2ers. In particular, I am interested in how these two age groups of L1-Korean L2ers perform on tasks investigating knowledge of the grammaticality and interpretation contrasts between VPE and Gapping.

[10] According to Schwartz (1997), regardless of whether the (exact) phenomenon of interest is present in the learner's native language, the L1 grammar always (initially) attempts to impose its analyses on the TL input. These analyses may then lead to an Interlanguage grammar that differs from the TL grammar. Furthermore, according to Schwartz & Sprouse (1996), failure on the part of the Interlanguage grammar to produce a well-formed parse of the TL input is what causes restructuring to occur.

3.4 First Language Processing Research on Gapping

There have been only a few native processing studies on Gapping in English (e.g., Carlson, 2001, 2002; Carlson et al., 2005; Hoeks et al., 2009; N. Kim et al., 2020; for Gapping in Dutch, see Kaan et al., 2013; for Gapping in German, see Claus, 2015; Hofmann, 2006; Streb et al., 2004). Carlson (2001) investigated the factors affecting the interpretive preferences of ambiguous Gapping sentences, as in (20), using a written questionnaire and an auditory comprehension task.

(20) Bill took chips to the party and Susan [e] to the game[a] / and [e] Susan to the game[b].

a. Subject reading (SR): 'Bill took chips to the party and Susan took chips to the game.'

b. Object reading (OR)[11]: 'Bill took chips to the party and Bill took Susan to the game.'

The SR in (20a) and the OR in (20b) have structurally different parses: SRs involve coordination of vP and ORs involve coordination of VP which is structurally smaller than vP is (for syntactic trees of these two parses, see (42) and (43) in §2.2).

In her written questionnaire, Carlson (2001) tested for the effects of (a) verb selection restrictions and (b) parallelism. For the former, she compared *bake*-type verbs vs. *amaze*-type verbs vs. *visit*-type verbs: As shown in (21), whereas *bake* usually takes non-human objects (e.g., (21a)) and *amaze* usually takes human objects (21c), *visit* can have either non-human objects (e.g., (21d)) or human objects (e.g., (21e)).

(21) a. Alice bakes cakes for tourists and brownies for her family. (Bake A)

b. Alice bakes cakes for tourists and Caroline for her family. (Bake B)

c. Dan amazed the judges with his talent and James with his musicality. (Amaze)

d. Josh visited the office during the vacation and Sarah during the week. (Visit A)

e. Josh visited Marjorie during the vacation and Sarah during the week. (Visit B)

(adapted from Carlson, 2001, p. 5, (7a)–(7e))

[11] Carlson (2001) used the terms 'Gapping reading' for the SR and 'non-Gapping reading' for the OR (for further discussion, see fn. 9 in Chapter 2).

She also manipulated the parallelism between arguments in the non-gapped clause and the gapped clause such that the first argument in the gapped clause, e.g., *Sarah* in (21d) and (21e), is parallel to the subject (e.g., *Josh* in (21d)), or to both the subject and the object in the non-gapped clause (e.g., *Josh* and *Marjorie* in (21e)).

In the written questionnaire, 68 participants were presented with seven *bake*-type sentences, five *amaze*-type sentences, and six *visit*-type sentences. They were then asked to (a) select the best paraphrase of each sentence, (b) rate on a 5-point scale how hard it was to comprehend the sentence, and (c) write what other meanings the sentence might have. The results showed that the sentences with SR-biased parallelism, like (21b) and (21d), received more SR responses than did the sentences with OR-biased parallelism, like (21a), or with weaker SR-biased parallelism, like (21c) and (21e). Moreover, effects of verb selection restrictions were found such that whereas *bake*-type sentences with SR-biased parallelism, as in (21b), yielded a strong preference for SR responses (81%) over OR responses (19%), *visit*-type sentences with intended SR-biased parallelism, as in (21d), still received more OR responses (60%) than SR responses (40%). The *amaze*-type sentences also received more OR responses (79%) relative to SR responses (21%). Furthermore, participants rated sentences as more difficult and provided more additional sentence meanings when they selected the SR responses vs. the OR responses, except for the unambiguous cases like (21b).

Carlson (2001) further explored whether prosody can contribute to the resolution of ambiguous Gapping sentences. In the auditory comprehension task, she manipulated the type, location, and range of pitch accents to make the ambiguous NP in the gapped clause contrast either with the subject, as in (22a), or with the object in the non-gapped clause, as in (22b).

(22) a. BOB insulted the guests during DINNER and SAM during the DANCE. (SR prosody)
b. Bob insulted the GUESTS during DINNER and SAM during the DANCE. (OR prosody)
(adapted from Carlson, 2001, p. 14, (11b) & (11c))

The auditory comprehension task asked participants to choose the best paraphrase of the sentence that they heard. The SR prosody led participants to choose SR responses significantly more often (44%) than the OR prosody did (28%). However, the SR prosody still obtained more OR responses. From this, Carlson (2001) concluded that parsers prefer the simpler structure

involving VP coordination (vs. vP coordination) except when it is semantically anomalous (e.g., (21b)), consistent with the Minimal Attachment account (e.g., Frazier, 1978; Frazier, 1987).

Because Carlson's (2001) study investigated which of two reading participants prefer (SR vs. OR) for Gapping sentences, as in (21), rather than whether or not participants accept each reading, a question remains as to whether native speakers do have knowledge of both readings. Furthermore, to examine the processing difficulty involved in a certain interpretation, she asked participants to rate the perceived difficulty, which is rather subjective. Study 2 in this dissertation addresses these issues (a) by using a picture-sentence matching task to test for knowledge of the two possible readings for Gapping and (b) by measuring reaction times for the judgments to investigate the processing difficulty associated with each reading.

N. Kim et al.'s (2020) recent eye-tracking-while-reading study examined (a) whether parsers posit verb gaps during online processing and (b) whether this processing is affected by connective sentence type (conjunction vs. adjunction) and/or parallelism (parallel vs. nonparallel). Their eye-tracking task had 32 items distributed across four conditions, as shown in (23). In the conjunct-parallel type sentences (e.g., (23a)), the NP-PP sequence in the second clause is (potentially) temporarily ambiguous. When encountering the string *the singer behind the stage* in (23a), parsers may build either a Gapping structure where the verb *hid* is gapped (i.e., *the singer [e] behind the stage*) or a non-Gapping structure where *the singer behind the stage* is analyzed as an NP subject to be followed by a predicate.

(23) a. The guitarist hid behind the curtain suddenly, and the singer behind the stage hid from the sneaky photographers. (conjunct-parallel)

b. The guitarist noticed his recording agent suddenly, and the singer behind the stage hid from the sneaky photographers. (conjunct-nonparallel)

c. The guitarist hid behind the curtain suddenly, whereas the singer behind the stage hid from the sneaky photographers. (adjunct-parallel)

d. The guitarist noticed his recording agent suddenly, whereas the singer behind the stage hid from the sneaky photographers. (adjunct-nonparallel)

(adapted from N. Kim et al., 2020, p. 789, (10a)–(10d))

The conjunct-nonparallel type, as in (23b), was created to see whether or not parsers would posit a verb gap when the conjunct clauses are not parallel. The adjunct counterparts of (23a) and (23b) were included, as in (23c) and (23d), neither of which can yield a Gapping structure. The three regions for analysis were (a) the verb region (e.g., *hid*), (b) the spill-over region 1 (e.g., *from the*), and the spill-over region 2 (e.g., *sneaky photographers*).

N. Kim et al. (2020) predicted that if parsers postulate a verb gap as soon as possible in a parallel conjunct clause, they should be surprised by the second verb (e.g., *hid* in (23a)) and there should thus be a slowdown only in this condition (because the others disallow Gapping). This prediction was borne out in their eye-tracking experiment with 52 English native speakers. Both the total fixation times at spill-over region 1 (e.g., *from the* in (23)) and the regression path durations at spill-over region 2 (e.g., *sneaky photographers* in (23)) were significantly longer in the conjunction-parallel condition than those in the other three conditions. This result indicates that parsers incrementally assign a Gapping analysis, positing the verb gap in real time, when it is grammatically licensed in conjunct clauses (vs. adjunct clauses) and when parallelism is observed.

More relevant to the processing research in this dissertation (Study 4) is Kaan et al.'s (2004) study, which used event related potentials (ERPs) to investigate the time course of identification and resolution of verb gaps. Their experimental sentences were constructed based on a plausibility manipulation paradigm (Garnsey, Tanenhaus, & Chapman, 1989). They manipulated the verb in the first clause so that the NP object in the gapped clause (e.g., *the hammer with the big head*) was either (a) a plausible object of the gapped verb, as in (24a), or (b) an implausible object of the verb, as in (24b). The two clauses in the Gapping sentences were always separated by a comma and the conjunction *and*. In most cases, the subjects of the two clauses were proper names. The critical regions were the first determiner and the head noun of the direct object in the gapped clause, e.g., *the hammer* in (24).

(24) a. Ron took the planks for the bookcase, and Bill the hammer with the big head.

b. *Ron sanded the planks for the bookcase, and Bill the hammer with the big head.

(Kaan et al., 2004, p. 591, Appendix A, (24))

Twenty-six English native speakers read 320 sentences in total, including 32 critical sentences; half of the trials were plausible (e.g., (24a)) and the other half were implausible (e.g., (24b)). All sentences were presented using word-by-word segmentation at a rate of 500 ms per word (the word appeared for 300 ms, followed by 200 ms of blank screen). After reading each sentence, they were asked to judge the sentence in terms of its syntactic/semantic acceptability.

The accuracy rates of the participants' acceptability judgments were 92% (*SD* = 11%) for the plausible sentences and 96% (*SD* = 5%) for the implausible sentences. Determiners after a gapped verb (e.g., *the* in *the hammer* in (24)) were accompanied by a negativity between 100 ms and 300 ms, followed by a positivity between 300 ms and 500 ms. This shows that the English native speakers recognized the verb gap as soon as they encountered the determiner. Importantly, in the implausible sentences (e.g., (24b)), relative to the plausible sentences (e.g., (24a)), the head noun region (e.g., *hammer* in (24)) displayed an N400 (300–500 ms), indicating a semantic anomaly effect. This N400 was immediately followed by a P600 (600–900 ms), which is associated with syntactic integration difficulty. These results indicate that the participants detected a verb gap and attempted to integrate the gapped verb with the object at the gap site.

However, Kaan et al. (2004) did not include baseline conditions to compare to the Gapping conditions (for discussion, see Kaan et al., 2013, pp. 308–309). Furthermore, element types following the gapped verb were not controlled, either; in some items, a complex NP object containing a PP modifier followed the verb gap, as in (24a), and in others, a direct object NP plus an indirect object NP followed it, as in (25).

(25) a. Tracy mailed the letter to George, and Julie the package to Lisa.

b. * Tracy wrote the letter to George, and Julie the package to Lisa.

(Kaan et al., 2004, p. 591, Appendix A, (23))

The present study addresses these issues (a) by including VPE as a baseline to compare to Gapping, and (b) by always using conjoined NP direct objects (e.g., *sandwiches and cake*) to control for the element type after the verb gap. In doing so, this study asks whether the successful verb-gap processing observed in Kaan et al.'s (2004) study with English native

speakers can be replicated and extended to L2 processing in an English self-paced reading study (see Chapter 7).

3.5 Summary of Chapter

This chapter reviewed previous literature on the L1 and L2 acquisition of VPE, the L2 acquisition of Gapping, and the native processing of Gapping (in English). First, L1 acquisition researchers have confirmed L1 children's early knowledge of VPE in English. Young L1 children accept only the possible interpretations of possessive pronouns unrealized in the ellipsis clause and reject the impossible ones (Foley et al., 2003; Matsuo, 2007). There have also been studies showing that children are sensitive to the structural parallelism in VPE vs. VPA (e.g., Matsuo & Duffield, 2001).

However, prior work on VPE in L2 English has yielded inconsistent results. Some studies have found that L2ers generally perform like native speakers (e.g., Al-Thubaiti, 2019; Hawkins, 2012; J. Kim, 2012, 2015). By contrast, L2ers in Duffield and Matsuo's (2009) study did not achieve target-like performance. I suggested that definitive conclusions cannot be made at this point just because even the very advanced L2ers, as a group, in the Al-Thubaiti and Hawkins studies did not show target-like performance on all VPE sentence types (viz. (8b), (13b)) and because we do not know whether the L2ers' performance in Duffield and Matsuo's study can be attributed to important factors, such as L2 proficiency or length of TL exposure.

As for Gapping, I reviewed two L2 adult studies, one by O'Grady (1999) and the other by Kanno (1999). L1-Japanese L2ers of English in the former study and L1-Mandarin L2ers of Japanese in the latter failed to accept the grammatical Gapping pattern in the TL, although they were able to reject the ungrammatical Gapping pattern. However, neither of these studies provided information about L2 proficiency or length of TL exposure, and so it remains unknown whether one or both of these factors might have contributed to the results.

Finally, I reviewed native English processing studies on Gapping. Using a written questionnaire and an auditory comprehension task, Carlson (2001) found significant effects of parallelism, verb selection restrictions, and prosody in the processing of Gapping. N. Kim et al.'s (2020) eye-tracking-while-reading study showed that parsers incrementally build a Gapping structure when the clausal conjuncts are parallel. In Kaan et al.'s (2004) ERP study, the English

native speakers identified a verb gap at the gap site and very quickly attempted to integrate the gapped verb with the following object.

I also brought up weaknesses of previous studies and identified a few important gaps in the research. I took these research gaps into consideration when designing my experiments.

3.6 Research Questions for the Present Dissertation

The primary motivation of this dissertation is the issue of learnability involved in the grammaticality and interpretation contrasts between VPE and Gapping in English (see §2.4), which has not been investigated either in L1 acquisition or in L2 acquisition. To address this issue, I conducted one corpus-based study and two acquisition studies. Study 1 applied natural language processing to (oral and written) corpora to investigate how (in)frequent VPE and Gapping are in the input to L1-English children and to L1-Korean L2ers of English (see Chapter 4). Study 2 and Study 3 tested L1-English children and (early and late) L1-Korean L2ers of English for knowledge of the previously discussed contrasts between VPE and Gapping: the grammaticality contrast via an acceptability judgment task (Study 2; see Chapter 5) and the interpretation contrast via a picture-sentence matching task (Study 3; see Chapter 6). To investigate any possible L2 proficiency effects in these two studies, a picture narration task (K.-S. Park, 2014) was administered to all the L2ers (as well as some of the L1 adults and L1 children), as an independent measure of proficiency. Furthermore, I explored the real-time processing of Gapping vs. VPE on the part of adult L1-Korean L2ers of English using a self-paced reading task in Study 4 (see Chapter 7); this study likewise tested for L2 proficiency effects among the L2ers by means of a cloze test (J. D. Brown, 1980). The research questions and methods for each of these four studies are laid out in Table 3.2.

Table 3.2
Overview of the Studies in the Present Dissertation

Study	Research Questions	Methods
1	• How (in)frequent are VPE and Gapping in input to L1-English children? • How (in)frequent are VPE and Gapping in input to L1-Korean L2ers of English?	Natural language processing analysis
2	• How early do L1-English-acquiring children know the contrast between licit vs. illicit VPE and Gapping in English? • Do early and late L1-Korean L2ers of English come to know the contrast between licit vs. illicit VPE and Gapping in English? What role does L2 proficiency play?	Acceptability judgment task, proficiency task
3	• How early do L1-English-acquiring children know the contrast between possible vs. impossible interpretations of VPE and Gapping in English? • Do early and late L1-Korean L2ers of English come to know the contrast between possible vs. impossible interpretations of VPE and Gapping in English? What role does L2 proficiency play?	Picture-sentence matching task, proficiency task
4	• Are adult L1-Korean L2ers of English able to recognize and resolve a verb gap in Gapping sentences in real time? What role does L2 proficiency play?	Self-paced reading task, proficiency task

I now turn to the natural language processing study, which investigated the incidence of VPE and Gapping in the input to L1-English children and in the input to L1-Korean L2ers of English.

CHAPTER IV
STUDY 1: NATURAL LANGUAGE PROCESSING ANALYSIS OF VP-ELLIPSIS INPUT AND GAPPING INPUT IN L1 AND L2 ACQUISITION

The main goal of Study1 is to identify how (in)frequent VP-Ellipsis (VPE) and Gapping are in the input for L1-English children and in the input for L1-Korean L2ers of English. In §4.1, I present my reanalysis of the VPE data reported in Bos and Spenader's (2011) study that used the *Wall Street Journal* corpus. Section 4.2 provides the research questions of this study. Section 4.3 introduces the (innovative) natural language processing analysis methods that were used to identify VPE and Gapping. Next, I report the results of my analyses of data selected from the Child Language Data Exchange System (CHILDES; MacWhinney, 2000) for the input to L1-English children (§4.4.1) as well as data obtained from various types of EFL (English as a Foreign Language) input given to L1-Korean L2ers of English (§4.4.2). Section 4.5 discusses the results of the study, focusing on learnability issues associated with the grammaticality and interpretation contrasts between VPE and Gapping. Finally, §4.6 concludes this chapter.

4.1 L1-English Corpora: Reanalysis of the VP-Ellipsis Data from Bos and Spenader (2011)

To date, only instances of VPE and closely related constructions (e.g., *Do So* Anaphora), but not Gapping, have been investigated using a corpus-based methodology (Bos & Spenader, 2011; Hardt, 1997; Nielsen, 2005). Most recently, Bos and Spenader (2011) analyzed the *Wall Street Journal* (WSJ) corpus (distributed by the Penn Treebank) for VPE. They found 487 instances of VPE (0.91%) out of 53,561 sentences in the WSJ corpus.

Since Bos and Spenader (2011) reported the frequency of VPE without considering the clause types in which it was used, I reanalyzed the data provided in their paper based on their annotations.[1] The reanalysis showed that among the 53,561 sentences in the corpus, 159 (0.30%) exhibited VPE in a conjunct clause (e.g., *Things were supposed to change when Vietnam's economic reforms gathered pace, and for a while they did*.) or in a separate sentence (e.g., *But early on, IBM offered its basic design to anybody wanting to copy it. Dozens of small companies did, swiftly establishing a standard operating system.*). Another 97 sentences (0.18%) had VPE

[1] Bos and Spenader's (2011) annotation method did not distinguish between a conjunct clause and a separate sentence (see pp. 473–481; 490–492).

in an adjunct clause (e.g., *On days when prices are tumbling, they must be willing to buy shares from sellers when no one else will.*) and 31 (0.06%) had VPE in antecedent contained deletion (e.g., *Maybe we recognize values the other guys don't, …*). One hundred and eighty-one instances of VPE (0.34%) appeared in comparative sentences (e.g., *Moreover, Japanese offices tend to use computers less efficiently than American offices do*.) or equative constructions (e.g., *He did not go as far as he could have in tax reductions; …*) and 4 instances (0.01%) occurred in tag questions (e.g., *But you knew that, didn't you?*). The remaining 15 instances of VPE (0.03%) appeared in a main clause preceded by an adjunct clause (e.g., *While the theme is compelling, the plot and characters are not.*).

Overall, VPE was infrequent both in conjunct clauses and in adjunct clauses in the native English corpus (see Bos & Spenader, 2011, p. 484, Table 4).

4.2 Research Questions

The following two research questions frame this study:

(a) How (in)frequent are VPE and Gapping in input to L1-English children?
(b) How (in)frequent are VPE and Gapping in input to L1-Korean L2ers of English?

4.3 Method

4.3.1 Data.

4.3.1.1 The input for L1-English children.

The corpus of input to L1-English children comprised utterances to children from mothers and fathers (etc.) provided in CHILDES (MacWhinney, 2000). These data were selected because they were directed to children in the age range of 3;0 to 6;0 and because they are considered a representative sample of the input that children receive (e.g., P. Li, 2000).[2] The selected data included Adam (3;0–5;2; R. Brown, 1973), Laura (3;0–5;10; Braunwald, 1997),

[2] Amy Schafer (personal communication, 17 April 2020) raised the possibility that the (in)frequency of VPE and Gapping may differ depending on the genre of the chosen corpus. For example, Montag (2019) found that children's picture books, relative to child-directed speech, contain a greater number of rare and complex sentence types, such as passives and relative clauses. Although it would be worthwhile to investigate genre effects in the frequency of VPE and Gapping, I believe that the input to L1-English children that I selected from CHILDES constitutes a suitable data source for this dissertation because it is a representative sample of the language exposure that children receive (e.g., P. Li, 2000) and it is unlikely that all caregivers read picture books to their children.

Ross (3;00–5;11; MacWhinney, 1991),[3] and Sarah (3;0–5;1; R. Brown, 1973). The native English input corpus thus contained 44,111 utterances for the 3-year-old sub-corpus, 28,447 for the 4-year-old sub-corpus, and 13,262 for the 5-year-old sub-corpus, making 85,820 utterances in total. The specific numbers of utterances and word tokens for each of the data sources are summarized in Table 4.1.

Table 4.1
Numbers of Utterances and Word Tokens for Each Sub-Corpus of the Native Input Corpus

Sub-corpus (age range)	CHILDES data	Number of utterances	Number of word tokens
3-year-old	Adam	11,381	42,302
	Laura	4,965	18,919
	Ross	12,778	49,791
	Sarah	14,987	50,164
	Sub-total	**44,111**	**161,176**
4-year-old	Adam	4,271	15,530
	Laura	3,501	14,902
	Ross	9,176	35,957
	Sarah	11,499	42,570
	Sub-total	**28,447**	**108,959**
5-year-old	Adam	317	1241
	Laura	10	45
	Ross	11,970	44,040
	Sarah	965	3,572
	Sub-total	**13,262**	**48,898**
TOTAL		**85,820**	**319,033**

4.3.1.2 The input for L1-Korean L2ers of English.

In order to create the corpus of input to L1-Korean L2ers of English, four types of data were collected, all in the L1-Korean context: (a) L1-Korean EFL teacher speech, (b) L1-English EFL teacher speech, (c) spoken input from EFL textbooks, and (d) written input from EFL textbooks. As for the EFL teacher talk data, 3,311 L1-Korean EFL teachers' utterances and 2,037

[3] The corpus of MacWhinney (1991) contains transcripts of recordings collected from his two sons. This study extracted only the transcripts targeting the older child Ross based on his age.

L1-English EFL teachers' utterances were transcribed from eight co-teaching[4] English classes, each at a different Korean elementary school. These classes were being observed by a representative from the Korean Office of Education for supervision and evaluation. The co-teaching classes analyzed in this study on average lasted 38 minutes and 27 seconds and consisted of about 25 Korean EFL learners (age range: 8–12) per class. The sub-corpus of the spoken input from EFL textbooks consisted of 13,900 utterances: 9,857 utterances were provided by Korean elementary school textbooks for 3rd to 6th graders and their supplementary materials, such as CD-ROMs (Ham et al., 2011, 2012; Lee et al., 2011, 2012); 1,899 were from Korean middle school textbooks for 7th to 9th graders (J. W. Kim, B. Ahn, Oh, S. Kim et al., 2013); and 2,144 were from Korean high school textbooks for 10th to 12th graders (J. W. Kim, B. Ahn, Oh, B. Kim et al., 2013; J. W. Kim, B. Ahn, Oh, Shin et al., 2013). The sub-corpus of the written input from EFL textbooks contained a total of 25,398 sentences, including 4,160 sentences from elementary school textbooks (Ham et al., 2011, 2012; Lee et al., 2011, 2012), 10,963 from middle school textbooks (J. W. Kim, B. Ahn, Oh, S. Kim et al., 2013), and 10,275 from high school textbooks (J. W. Kim, B. Ahn, Oh, B. Kim et al., 2013; J. W. Kim, B. Ahn, Oh, Shin et al., 2013). Detailed information about each of these sub-corpora is provided in Table 4.2.

[4] It has been a regular practice in Korea that an L1-Korean EFL instructor and an L1-English EFL instructor teach English together (M. Kim, 2010).

Table 4.2
Numbers of Utterances/Sentences and Word Tokens for Each Sub-Corpus of the EFL Input Corpus

Sub-corpus (EFL input type)	Data	Number of utterances	Number of word tokens
L1-Korean EFL teacher speech	L1-KOR1	398	1,329
	L1-KOR2	470	2,245
	L1-KOR3	333	1,520
	L1-KOR4	594	1,736
	L1-KOR5	420	1,551
	L1-KOR6	230	713
	L1-KOR7	486	1,382
	L1-KOR8	380	1,347
	Sub-total	**3,311**	**11,823**
L1-English EFL teacher speech	L1-ENG1	348	2,109
	L1-ENG2	270	1,112
	L1-ENG3	264	1,073
	L1-ENG4	214	654
	L1-ENG5	223	866
	L1-ENG6	124	417
	L1-ENG7	380	1,120
	L1-ENG8	214	763
	Sub-total	**2,037**	**8,114**
Spoken input from EFL textbooks	Elementary	9,857	30,886
	Middle	1,899	9,888
	High	2,144	14,429
	Sub-total	**13,900**	**55,203**
Written input from EFL textbooks	Elementary	4,160	12,451
	Middle	10,963	53,188
	High	10,275	66,509
	Sub-total	**25,398**	**132,148**
TOTAL		**44,646**	**207,288**

Note. In order to preserve confidentiality, the data names for the teacher talk were coded using combinations of characters and numbers, such as L1-KOR1 (L1-Korean EFL teacher #1) and L1-ENG1 (L1-English EFL teacher #1).

4.3.2 Natural language processing analysis.

The data were analyzed in four stages in Python 3.7.1 (Python Software Foundation, 2018). In the first stage, the text was tokenized into sentences using the Natural Language Toolkit (NLTK; Bird, Klein, & Loper, 2009). The second stage depended on the construction. For VPE, a sentence was split into clauses and then a search was run for clauses containing an auxiliary verb (e.g., *be*, *do*, *have*), a modal verb, negation (*not*), or a *to*-infinitive, all of which are

potential identifiers for this construction. If the word in question was not followed by a verb, the clause was extracted as a VPE candidate. This step was performed using the spaCy library (Honnibal & Montani, to appear), which consists of a suite of tools for part-of-speech tagging and dependency parsing. In the case of Gapping, for the second stage, any clause that lacks a verb was extracted as a Gapping candidate using the Benepar library (Kitaev & Klein, 2018). The full Python script for the first and second stages can be found at https://github.com/Haerim-Hwang/NLP_Python/tree/master/Identify_VPE_Gapping. Each sentence that passed the second stage was manually checked for instances of VPE and Gapping in the third stage. Lastly, VPE and Gapping instances were classified into different categories. VPE instances were divided into eight categories: (a) VPE in a sentence that is separate from the antecedent, henceforth VPE in a separate sentence (e.g., *I read books. Mommy does, too.*); (b) VPE in a conjunct clause (e.g., *I read books, and mom does, too.*); (c) VPE in an adjunct clause (e.g., *I read books because mom does.*); (d) VPE in a complement clause (e.g., *Mom thinks that I love books, but dad knows that I don't.*); (e) VPE in antecedent contained deletion (ACD) (e.g., *I read every book that mom did.*); (f) VPE in a comparative construction (e.g., *I love books more than mom does.*) or an equative construction (e.g., *I love books as much as mom does.*); (g) VPE in a tag question (e.g., *Mommy is reading a book, isn't she?*); (h) miscellaneous[5]. Gapping sentences were sorted into two categories: (a) subject reading Gapping (e.g., *I sat on the chair and mom on the sofa.*) and (b) object reading Gapping (e.g., *I put the flower on the table and the books on the desk.*).

4.4 Results

4.4.1 VP-Ellipsis and Gapping in input to L1-English children.

This section reports the incidence of VPE and Gapping in the whole native input corpus, rather than in each of the sub-corpora, because there were not any considerable differences between the frequency distributions for the sub-corpora (see Table 4.3).

[5] This category consisted of instances which could not be grouped into any of the other categories (e.g., [Laura_041000] *When we're all done, you can.*).

Table 4.3

Frequencies of VPE and Gapping Instances in Each Sub-Corpus of the Native Input Corpus

		VPE									Gapping		
Sub-corpus (age range)	**CHILDES data**	separate[a]	conjunct[b]	adjunct[c]	complement[d]	ACD[e]	comparative /equative[f]	tag[g]	etc.[h]	**Total**	SR[i]	OR[j]	**Total**
3-year-old	Adam (11,381)	285	1	14	19	-	6	9	2	**336 (2.95%)**	-	-	**-**
	Laura (4,965)	102	5	6	12	-	-	49	2	**176 (3.54%)**	-	-	**-**
	Ross (12,778)	262	4	6	7	4	2	15	-	**300 (2.35%)**	-	1	**-**
	Sarah (14,987)	249	9	3	-	-	4	15	2	**282 (1.88%)**	1	1	**-**
	Sub-total (44,111)	**898 (2.04%)**	**19 (0.04%)**	**29 (0.07%)**	**38 (0.09%)**	**4 (0.01%)**	**12 (0.03%)**	**88 (0.20%)**	**6 (0.01%)**	**1,094 (2.48%)**	**1 (0.00%)**	**2 (0.00%)**	**3 (0.01%)**
4-year-old	Adam (4,271)	111	-	2	3	-	3	-	-	**119 (2.79%)**	-	1	**-**
	Laura (3,501)	53	1	2	2	-	2	18	4	**82 (2.34%)**	-	-	**-**
	Ross (9,176)	155	6	14	6	1	9	5	-	**196 (2.14%)**	-	-	**-**
	Sarah (11,499)	244	4	1	-	1	3	3	-	**256 (2.23%)**	-	-	**-**
	Sub-total (28,447)	**563 (1.98%)**	**11 (0.04%)**	**19 (0.07%)**	**11 (0.04%)**	**2 (0.01%)**	**17 (0.06%)**	**26 (0.09%)**	**4 (0.01%)**	**653 (2.30%)**	-	**1 (0.00%)**	**1 (0.00%)**
5-year-old	Adam (317)	6	-	-	-	-	-	-	-	**6 (1.89%)**	-	-	**-**
	Laura (10)	-	-	-	-	-	-	-	-	-	-	-	**-**
	Ross (11,970)	183	5	9	6	-	5	18	1	**227 (1.90%)**	-	-	**-**
	Sarah (965)	10	-	-	1	-	1	-	-	**12 (1.24%)**	-	-	**-**
	Sub-total (13,262)	**199 (1.50%)**	**5 (0.04%)**	**9 (0.07%)**	**7 (0.05%)**	**-**	**6 (0.05%)**	**18 (0.14%)**	**1 (0.01%)**	**245 (1.85%)**	-	-	**-**
TOTAL (85,820)		**1,660 (1.93%)**	**35 (0.04%)**	**57 (0.07%)**	**56 (0.07%)**	**6 (0.01%)**	**35 (0.04%)**	**132 (0.15%)**	**11 (0.01%)**	**1,992 (2.32%)**	**1 (0.00%)**	**3 (0.00%)**	**4 (0.00%)**

Notes. The eight categories for VPE are: [a]VPE in a separate sentence, [b]VPE in a conjunct clause, [c]VPE in an adjunct clause, [d]VPE in a complement clause, [e]VPE in antecedent contained deletion (ACD), [f]VPE in a comparative/equative construction, [g]VPE in a tag question, and [h]miscellaneous. The two categories for Gapping are [i]subject reading (SR) Gapping and [j]object reading (OR) Gapping.

Gapping was rare in the parental input given to L1-English children. There were only three occurrences (0.003%) of object reading Gapping (e.g., [Adam_041002] *You may give Robin a banana and Ursula a banana if you like to.*; [Ross_030105d] *I'll give you one and one for Mark.*; [Sarah_031009] *They brought Mama ice-cream and Auntie back ice-cream.*) and one (0.001%) of subject reading Gapping (e.g., [Sarah_030328] *It has one line going across and one line down.*).

There were many more utterances containing VPE than Gapping. A total of 1,992 utterances involved VPE (2.32%). Interestingly, the frequency of VPE decreased slightly as children's ages increased: 1,094 instances (2.48%) in the 3-year-old corpus, 653 instances (2.30%) in the 4-year-old corpus, and 245 instances (1.85%) in the 5-year-old corpus. This pattern resulted from a steady decline in the frequency of the most common type of VPE in the corpora, namely VPE in a separate sentence. There were 898 such instances (2.04%) in the 3-year-old corpus, 563 (1.98%) in the 4-year-old corpus, and 199 (1.50%) in the 5-year-old corpus, making a total of 1,660 instances in the whole native input corpus (1.93%). These instances of VPE came mostly in the form of questions (e.g., (1)) and responses to previous utterances (e.g., (2)).

(1) Child: Pretty soon I have to have… be on a diet.
Mother: You will?

(from the CHILDES data: Laura_040211b)

(2) Child: Let's go to California now.
Father: No, you can't.

(from the CHILDES data: Ross_031114)

Interestingly, there were 16 cases of VPE (0.02%), which occurred in adjunct clauses that stand alone without any main clause, as shown in (3).

(3) Child: Why he plays two of them?
Mother: Because he likes to.

(from the CHILDES data: Adam_030501)

Including such cases, 57 occurrences of VPE (0.07%) were in adjunct clauses (e.g., [Adam_030418] *But you may be a doctor when you grow up if you'd like to.*). Thirty-five occurrences of VPE (0.04%) were in conjunct clauses (e.g., [Sarah_031016] *You could say it and I couldn't.*). VPE also appeared in complement clauses 56 times (0.07%); in each of these cases, the antecedent was in the previous utterance, as in (4).

(4) Child: I won't spill this on the floor.
Mother: I know you won't.

(from the CHILDES data: Adam_040624)

Six cases of VPE (0.01%) occurred in ACD (e.g., [Ross_030522] *Everything was closed so I couldn't buy any stuff that I wanted to.*), 35 cases (0.04%) in comparative or equative constructions (e.g., [Laura_040718c] *A harp is even bigger than I am.*), and 132 cases (0.15%) in tag questions (e.g., [Adam_030826] *It looks more like a big violin, doesn't it?*). The remaining 11 cases (0.01%), classified as miscellaneous, included VPE appearing in a main clause preceded by an adjunct clause (e.g., [Laura_041000] *When we're all done, you can.*) and the combination of two instances of VPE, as shown in (5).

(5) Child: You can't at my house, okay?
Mother: I can or I can't?

(from the CHILDES data: Laura_030627b)

4.4.2 VP-Ellipsis and Gapping in input to L1-Korean L2 learners of English.

The four types of data collected for the EFL corpus exhibited similar frequency distributions of the categories made for VPE and Gapping, although the spoken input from EFL textbooks showed more occurrences of VPE in separate utterances than the other types of EFL input, as shown in Table 4.4. I therefore report the results of the analysis made on the whole EFL corpus.

Table 4.4

Frequencies of VPE and Gapping Instances in Each Sub-Corpus of the EFL Input Corpus

		VPE									Gapping		
Sub-corpus (EFL input type)		separate[a]	conjunct[b]	adjunct[c]	complement[d]	ACD[e]	comparative /equative[f]	tag[g]	etc.[h]	**Total**	SR[i]	OR[j]	**Total**
L1-Korean EFL teacher speech (3,311)		**14 (0.42%)**	-	-	-	-	**1 (0.03%)**	-	-	**15 (0.45%)**	-	-	-
L1-English EFL teacher speech (2,037)		**12 (0.59%)**	-	-	-	-	-	**1 (0.05%)**	-	**13 (0.64%)**	**1 (0.05%)**	-	**1 (0.05%)**
Spoken input from EFL textbooks (13,900)	Elementary (9,857)	309 (3.13%)	-	-	-	-	-	-	-	**309 (3.13%)**	-	-	-
	Middle (1,899)	35 (1.84%)	-	-	-	-	-	1 (0.05%)	-	**36 (1.90%)**	-	-	-
	High (2,144)	15 (0.70%)	-	-	-	-	-	9 (0.42%)	-	**24 (1.12%)**	-	-	-
	Sub-total	**359 (2.58%)**	-	-	-	-	-	**10 (0.07%)**	-	**369 (2.65%)**	-	-	-
Written input from EFL textbooks (25,398)	Elementary (4,160)	68 (1.63%)	-	-	-	-	-	-	-	**68 (1.63%)**	-	-	-
	Middle (10,963)	90 (0.82%)	-	-	-	-	3 (0.03%)	14 (0.13%)	-	**107 (0.98%)**	-	-	-
	High (10,275)	18 (0.18%)	-	1 (0.01%)	-	-	3 (0.03%)	7 (0.07%)	-	**29 (0.28%)**	1 (0.01%)	-	1 (0.01%)
	Sub-total	**176 (0.69%)**	-	**1 (0.00%)**	-	-	**6 (0.02%)**	**21 (0.08%)**	-	**204 (0.80%)**	**1 (0.00%)**	-	**1 (0.00%)**
TOTAL (44,656)		**561 (1.26%)**	-	**1 (0.00%)**	-	-	**7 (0.02%)**	**32 (0.07%)**	-	**601 (1.35%)**	**2 (0.00%)**	-	**2 (0.00%)**

Notes. The eight categories for VPE are: [a]VPE in a separate sentence, [b]VPE in a conjunct clause, [c]VPE in an adjunct clause, [d]VPE in a complement clause, [e]VPE in antecedent contained deletion (ACD), [f]VPE in a comparative/equative construction, [g]VPE in a tag question, and [h]miscellaneous. The two categories for Gapping are [i]subject reading (SR) Gapping and [j]object reading (OR) Gapping.

There were only two occurrences of Gapping (0.004%) in the EFL input, one of which appeared in L1-English EFL teacher talk (e.g., *Can I have the boys make one line here and our girls the other line here?*) and the other one in the written input from high school EFL textbooks (e.g., *Both groups took a walk for about 50 minutes, but one group took a walk in the woods and the other in the city.*). Both occurrences had a subject reading.

By contrast, VPE was observed relatively frequently in separate utterances in cross-dialogue, as was the case in the native input corpus to children. All of these 561 cases (1.26%) happened to be either questions, as illustrated in (6), or answers to the previous utterance made by the other speaker in the interaction context, as in (7).

(6) Q: I bought it online.
A: Did you?

(from the written input in high school EFL textbooks)

(7) Q: Is Chris studying?
A: No, he isn't.

(from the spoken input in elementary school EFL textbooks)

It should be pointed out that the rather high frequency of this type came from the spoken input from EFL textbooks, which had 359 such occurrences (2.58%); these were mostly from the elementary school textbooks (309 cases; 3.13%). Consistent with the native English input corpus, the instances of this VPE type decreased in the spoken input from EFL textbooks, as school level increased: e.g., elementary school (309 cases; 3.13%), middle school (35 cases; 1.84%), and high school (15 cases; 0.70%). (The same pattern was also observed in the written input from EFL textbooks.) There were fewer occurrences of this VPE type in the L1-Korean EFL teacher speech (14 cases; 0.42%), the L1-English EFL teacher speech (12 cases; 0.59%), and the written input from EFL textbooks (176 cases; 0.69%). These cases were all main clause questions and answers. There were no exchanges like the one shown in (3).

The other types of VPE appeared very infrequently. There was one instance (0.002%) in an adjunct clause (e.g., [Written input from high school EFL textbooks] *It helps students study whenever and wherever they want to.*), seven instances (0.02%) in comparative/equative

constructions (e.g., [Written input from middle school EFL textbooks] *They are still friends, but their friendship is not as strong as it was.*), and 32 instances (0.07%) in tag questions (e.g., [Written input from high school EFL textbooks] *This looks like a regular glue stick, doesn't it?*). Interestingly, VPE never appeared in conjunct clauses in the EFL input corpus.

4.5 Discussion

The input to L1-English children and the input to L1-Korean L2ers of English contained very few occurrences of Gapping, and there were not many cases of VPE in conjunct clauses, adjunct clauses, complement clauses, ACDs, comparative/equative constructions, or tag questions, either. Nevertheless, there were relatively many instances of VPE occurring in separate sentences from their antecedents in both the native input to children and the spoken input from EFL textbooks.

However, it is highly doubtful that such input can successfully lead both L1-English children and L1-Korean L2ers of English to acquire the grammaticality contrast between VPE and Gapping in English—and specifically the fact that Gapping is ungrammatical in adjunct clauses. If it were the case that these learner groups use analogy either (a) between Gapping in conjunct clauses and Gapping in adjunct clauses or (b) between VPE in adjunct clauses and Gapping in adjunct clauses, then we would expect them to incorrectly allow Gapping in adjunct clauses (see §2.4). For the L1-Korean L2ers, the implicit knowledge that VPE in adjunct clauses is grammatical cannot come from their L1 grammar, either. As discussed in §2.1, Korean does not have a true equivalent to VPE although it has three false analogues of VPE. Pseudo-VPE cannot occur in adjunct clauses at all. Argument Ellipsis (AE) can appear in adjunct clauses only when its subject in the ellipsis clause has the nominative case marker (vs. *-to* 'also'), and *Kulay* 'Do So' Anaphora can occur in adjunct clauses only when it follows its antecedent that is in a separate sentence or main clause. However, there is no reason to believe that L1-Korean L2ers of English initially assume that VPE has the same distributional properties as AE and *Kulay* 'Do So' Anaphora. In addition, the grammaticality contrast at issue cannot come from the L2ers' classroom instruction since neither VPE nor Gapping has been targeted in the Korean National Curriculum (Ministry of Education, 2015).

As for the interpretation contrast between VPE and Gapping in English, this too cannot be acquired solely from the input that L1-English children and L1-Korean L2ers of English

receive. Specifically, implicit knowledge of the impossibility of the object reading for VPE is not derivable from the input alone. Nor is it possible for the L1-Korean L2ers to become sensitive to the fact that English does not permit the object reading for VPE based on their L1 grammar: Korean does not have VPE and even the false Korean analogues to VPE (AE, *Kulay* 'Do So' Anaphora, Pseudo-VPE) all permit both a subject reading and an object reading (see §2.1). These L2ers' classroom instruction is not helpful, either, for them to acquire the interpretation contrast between VPE and Gapping because, again, neither is a topic of instruction in Korean EFL classrooms (see §2.4.2).

In summary, the grammaticality and interpretation contrasts between VPE and Gapping in English raise learnability problems for L1-English children and L1-Korean L2ers of English. Therefore, if these groups demonstrate knowledge of these contrasts, it would therefore indicate that they are able to overcome the learnability problems.

4.6 Conclusion

This natural language processing study revealed hardly any instances of VPE in adjunct clauses or of Gapping at all in the input corpora to L1-English children and to L1-Korean L2ers of English. This suggests that input alone cannot derive the grammaticality and interpretation contrasts between VPE and Gapping in both L1 acquisition of English and L1-Korean speakers' L2 acquisition of English.

CHAPTER V
STUDY 2: L1 AND L2 ACQUISITION OF THE GRAMMATICALITY CONTRAST BETWEEN VP-ELLIPSIS AND GAPPING IN ENGLISH

This chapter reports a study that examined whether L1-English children and L1-Korean early and late L2ers of English (come to) have knowledge of the grammaticality contrast between VP-Ellipsis (VPE) and Gapping. Section 5.1 lists the research questions of this study. In §5.2, I present the method of the acceptability judgment task, the principal task employed by this study to test for knowledge of the target contrast. Section 5.3 reports the judgment data from this task. The subsequent section (§5.4) discuss the results of the study by also examining reaction time data. Section 5.5 concludes this chapter.

5.1 Research Questions

This study addresses two main research questions:

(a) How early do L1-English-acquiring children know the contrast between licit vs. illicit VPE and Gapping in English?
(b) Do early and late L1-Korean L2ers of English come to know the contrast between licit vs. illicit VPE and Gapping in English? What role does L2 proficiency play?

5.2 Method

5.2.1 Participants.

Four participant groups took part in this study: (a) native English-speaking adult controls ("L1 adults," n = 70), (b) native English-speaking children ("L1 children," n = 46), (c) L1-Korean early L2ers of English ("early L2ers," n = 43), and (d) L1-Korean late L2ers of English ("late L2ers," n = 31). The L1 adults and L1 children were recruited in Honolulu, Hawai‘i. The L1 children consisted of 3-year-olds (n = 2), 4-year-olds (n = 3), 5-year-olds (n = 13), 6-year-olds (n = 19), and 7-year-olds (n = 9). All L2ers were recruited in Seoul, Korea.

This dissertation follows K. Kim (2014) in sorting participants into groups based on the age at which they were first exposed to English: The L1 children had their first exposure at birth, the early L2ers had their first exposure between ages 4 and 6 (Schwartz, 2004), and the late

L2ers received their first exposure between ages 8 and 12. L1-Korean children who had been exposed to English before age 4 were not included in the study because previous research has shown that children master most of their L1 grammar by age 4 (Guasti, 2002), and therefore L1-Korean children who start learning English before this age can arguably be considered simultaneous bilinguals. The cutoff age between the early and late L2ers was set at 8 because children who begin L2 acquisition before age 8 have been shown to possess target-like knowledge for a variety of morpho-syntactic phenomena (e.g., Johnson & Newport, 1989), and thus it can reasonably be expected that there will be a qualitative difference between participants who start learning English above and below this age. In order to make a clear distinction between the early and late L2ers, I excluded learners who had their first exposure to English at age 7.

A picture narration task (PNT) was administered to all the L2 participants (and to 32 of the L1 adults and 32 of the L1 children) to measure their English proficiency, following K.-S. Park (2014), Unsworth (2005), and Whong-Barr and Schwartz (2002). This independent proficiency test was regarded as appropriate for both adults and children because it did not require unfamiliar vocabulary items or complex syntax. It is therefore reasonable to use the English proficiency scores obtained from this task as a guideline to compare the performance of the early L2ers vs. the late L2ers on the tasks testing for the VPE-Gapping contrasts with respect to grammaticality (Study 2) and interpretation (Study 3). The PNT consisted of three sets of four pictures depicting everyday activities (i.e., a morning routine, fighting between friends, and nighttime parenting) in sequential order (see K.-S. Park, 2014, p. 146). For each set, participants were asked to tell a story in English based on a sequence of pictures presented via PowerPoint. I followed K.-S. Park and Unsworth for the computation of English proficiency scores. Unsworth operationalized L2 proficiency as "the ability to produce lexically, morphologically, and syntactically complex and accurate utterances in the target language" (pp. 154–155). Thus, the following three measures were used to analyze the elicited production data from the PNT: (a) morpho-syntactic complexity (total number of verbs divided by the total number of T-units[1]), (b) lexical complexity (average of the type-token ratio for every moving text sequence of 15 consecutive words; Covington & McFall, 2010), and (c) morphological/syntactic/lexical

[1] A T-unit is defined as "one main clause plus whatever subordinate clauses happen to be attached to or embedded within it" (Hunt, 1966, p. 735).

accuracy (total number of error-free T-units divided by the total number of T-units). Appendix A provides details of the method and results (both by group and by individual) of the PNT.

I wanted an inclusion/exclusion criterion that would retain as many participants as possible who were able to reject ungrammatical sentences in a non-random fashion. The inclusion criterion was therefore set as *able to reject at least two out of 10 impossible* wanna-*contraction items* (see, e.g., (3b), (3c) below), which were one type of filler in the main acceptability judgment task (see §5.2.2.1).[2] Applying this criterion excluded 13 L1 children, 14 early L2ers, and one late L2er. Two additional early L2ers were excluded because they did not produce any sentence-level utterances in the PNT, which made it impossible to evaluate their level of English proficiency. The aggregated background information for the remaining participants in each group is given in Table 5.1.

Table 5.1
Background Information for Participants in Study 2

	Age at testing	**Age of English onset**	**Length of residence in an English-speaking country in months**
L1 Adults (*n* = 70)	23.27 (*SD* = 5.19; range = 18–49)	N/A	N/A
L1 Children (*n* = 33)	5.76 (*SD* = 1.03; range = 3–7)	N/A	N/A
Early L2ers (*n* = 27)	8.52 (*SD* = 1.63; range = 5–12)	4.96 (*SD* = 0.76; range = 4–6)	1.15 (*SD* = 3.46; range = 0–13)
Late L2ers (*n* = 30)	23.03 (*SD* = 2.92; range = 18–30)	8.83 (*SD* = 1.09; range = 8–12)	2.83 (*SD* = 5.98; range = 0–24)

The early L2ers older than age 8 and all the late L2ers were under the national English curriculum in Korea (Ministry of Education, 2015) at the time of testing. This curriculum aims to develop students' communicative competence in listening, speaking, reading, and writing. In elementary school, 3rd and 4th graders learn English for 2 lesson hours (80 minutes) per week; 5th and 6th graders learn English for 3 lesson hours (120 minutes) per week. For 7th to 9th

[2] I am grateful to Bonnie D. Schwartz (personal communication, 6 May 2020) for helpful discussion concerning this criterion.

graders in middle school, approximately 3–4 lesson hours (135–180 minutes) per week are assigned to English classes, and for 10th to 12th graders in high school, 4.5–5 lesson hours (225–250 minutes) per week are allocated to English classes. Class size is approximately 25 students.

A unique identifier was assigned to each participant to ensure anonymity. For example, the identifiers L1A_01, L1C_01, EL2_01, and LL2_01 were given to the first participants tested and included for analysis in the L1 adult group, the L1 child group, the early L2er group, and the late L2er group, respectively.

5.2.2 Acceptability judgment task.

5.2.2.1 Materials.

The stimuli comprised 24 critical sentences and 44 filler sentences. All these sentences were 6 to 8 words long and consisted of simple and highly frequent words to minimize the cognitive burden on participants. The 24 critical sentences were distributed in a 2 × 2 Latin square design crossing the factors *Construction* (VPE; Gapping) and *Clause* (Conjunct; Adjunct), as shown in (1), which resulted in four running lists. Conjunct clauses were introduced with *and* and adjunct clauses were introduced with *because*. A full list of all experimental items (including the fillers) is provided in Appendix B.

(1) Critical sentences

a. VPE in a conjunct clause (VPE-C; $k = 6$)
   Sara made pizza, and Kelly did too.

b. VPE in an adjunct clause (VPE-A; $k = 6$)
   Sara made pizza because Kelly did.

c. Gapping in a conjunct clause (Gapping-C; $k = 6$)
   Sara made pizza, and Kelly pasta.

d. Gapping in an adjunct clause (Gapping-A; $k = 6$)
   * Sara made pizza because Kelly pasta.

There were 44 bi-clausal fillers consisting of 16 grammatical sentences and 28 ungrammatical sentences such that the number of grammatical and ungrammatical sentences was

balanced across all the stimuli. An example of each type of filler sentence is provided in (2), (3), (4), and (5).[3] To test for L1 transfer, I included three filler sentences involving the ungrammatical backward Gapping pattern (e.g., (5)), which would be grammatical (at some level of abstraction) in Korean. The results of each filler type for all four participant groups are provided in Appendix C.

(2) Fillers: null vs. overt argument

a. No null argument ($k = 6$)

The ball was dirty, but I loved it.

b. Null subject ($k = 6$)

* The ball was dirty, but loved it.

c. Null object ($k = 6$)

* The ball was dirty, but I loved.

(3) Fillers: *wanna* contraction

a. If + No gap ($k = 5$)

I wonder if you wanna work.

b. If + Gap ($k = 5$)

* I wonder if you wanna work with.

c. Who + No gap ($k = 5$)

* I wonder who you wanna work.

d. Who + Gap ($k = 5$)

I wonder who you wanna work with.

[3] The fillers were originally designed to test L1-Korean L2ers' knowledge of the obligatoriness of arguments in English (e.g., (2)), *wanna* contraction (e.g., (3)), and 3sg present [–s] subject–verb agreement (e.g., (4)). Admittedly, these fillers were not good choices for screening out L1-English children not able to reject ungrammaticality, i.e., with a strong *yes*-bias, since young L1-English kids have been found to allow null subjects and null objects (e.g., Hyams, 1986, 2011; Jaeggli & Hyams, 1988; Orfitelli & Hyams, 2012), *wanna* in illicit contexts (Zukowski & Larsen, 2011), and stem forms with 3sg subjects (R. Brown, 1973; Johnson, de Villiers, & Seymour, 2005; Legendre, Culbertson, Zaroukian, Hsin, Barrière, & Nazzi, 2014).

(4) Fillers: missing 3sg present [–s] subject–verb agreement (k = 3)

* Tony says that his mom want a car.

(5) Fillers: ungrammatical backward gapping pattern (k = 3)

* Ryan the chair, and I liked the desk.

For each item, there was an audio stimulus as well as its corresponding written sentence. The audio stimuli for the ungrammatical sentences were constructed with cross-splicing techniques using Praat (Boersma & Weenink, 2017). For example, the audio stimulus for (1d) was created by slicing out the first part of (1b), i.e., *Sara made pizza because*, and the last part of (1c), i.e., *Kelly pasta*, and then combining these two segments. All splicing was done at the nearest zero-crossing points. Caution was taken to control for the pause length between the two clauses (e.g., [1] in (6)) and at the gapped region in the case of Gapping (e.g., [2] in (6)).

(6) Sara made pizza, [1] and Kelly [2] pasta.

Also, the location and choice of pitch accents and the pitch range were checked for consistency within each condition. The prosody of a sample stimulus for each critical condition is presented with a pitch track in Appendix D.

5.2.2.2 Procedure.

All participants were tested individually in a quiet room. They began by filling out a language background questionnaire (see Appendix E); in the case of the child participants, it was their parents who filled out the questionnaire.

The acceptability judgment task (AJT) was designed and administered in PsychoPy (Peirce, 2017). Each trial began with an audio stimulus, which was presented twice with a one-second interval in between. At the same time, the corresponding written sentence was displayed on the computer screen. At the offset of the audio stimulus, a 4-point "smiley face" scale (see Figure 5.1) popped up on the screen.

Figure 5.1
Four-Point "Smiley Face" Scale for the Acceptability Judgment Task

Note. Adapted from Ambridge, Pine, Rowland, and Young (2008).

The smiley faces were described as follows during the oral instructions for the task: 'very bad/definitely impossible,' 'bad/impossible,' 'good/possible,' and 'very good/definitely possible.' An additional 'I don't know' option, presented in the form of a question mark, was made available to participants in case they could not rate the sentence for some reason. Participants pressed these images, which had been attached to buttons on the keyboard, to provide their judgments. PsychoPy recorded the participants' judgments as well as their reaction times (RTs; measured from the offset of the audio stimulus). Judgments served as my primary data. RTs were collected as supplementary data, which were expected to help better understand more nuanced aspects of participants' judgments.

After completing four practice sentences, participants proceeded to the experimental session, which was divided into two blocks with a break of approximately 5–10 minutes in between to prevent participants, particularly the children, from losing interest. Each block consisted of 12 critical sentences and 22 fillers presented in a pseudo-random order such that no two stimuli from the same condition were presented consecutively. Completing this task took approximately 20 minutes including the break.

Next came the picture-sentence matching task (see Study 3, Chapter 6) and then the PNT (see Appendix A) with a break of 3–10 minutes between them. These two tasks were administered to all L2 participants but to only 32 of the L1 adults and 32 of the L1 children; the screening procedure described in §5.2.1 finally left me with 32 L1 adults, 24 L1 children, 27 early L2ers, and 30 late L2ers.The picture-sentence matching task took about 20 minutes and the PNT took about 10 minutes.

5.2.3 Data analysis.

The acceptability judgment data consisted of both sentence ratings and RTs. Prior to statistical analysis, the 'I don't know' judgments were excluded, which led to the removal of 0.15% of the L1 adult data, 3.48% of the L1 child data, 5.23% of the early L2 data, and 0.20% of the late L2 data.

All sentence ratings were transformed to binary format for analysis. The upper two response options on the scale were recoded as "1" (accept), and the lower two response options were recoded as "0" (reject). This transformation was regarded as valid because of the way the smiley faces were described in the instructions ('very bad/definitely impossible,' 'bad/impossible,' 'good/possible,' 'very good/definitely possible'). Next, a logistic mixed-effects model was fitted to the binary values for each group, with *Construction* and *Clause* as binary fixed effects (contrast-coded with [−.5, .5]) and *participant* and *item* as random effects. The L1 child data were further examined for a potential effect of *Age* by constructing an additional logistic mixed-effects model in which *Age* was included as a continuous fixed effect in addition to the fixed effects *Construction* and *Clause* (contrast-coded with [−.5, .5]) and the random effects *participant* and *item*. *Age* was computed as the amount of time intervening between the participant's date of birth and date of testing. For the L2 data, the effect of *Proficiency* was investigated using a separate mixed-effects model with the fixed effect *Proficiency* included in addition to the fixed effects *Construction* and *Clause* (contrast-coded with [−.5, .5]) and the random effects *participant* and *item*. Given that proficiency is gradient in nature (e.g., Newman, Tremblay, Nichols, Neville, & Ullman, 2012; Tanner, Inoue, & Osterhout, 2014; van Hell & Tanner, 2012), proficiency scores were treated as continuous in the regression model.

I also checked for proficiency effects in the L2 data using descriptive statistics by dividing the L2ers into three groups based on their *z*-transformed final proficiency scores (see K. Kim, 2014; Unsworth, 2005): 'Lower' (z-score below −0.5), 'Medium' (*z*-score between −0.5 and 0.5), and 'Higher' (*z*-score above 0.5). The *z*-transformation procedure was performed on the combined data from the early and late L2ers so that comparisons could be made between early and late L2 acquisition of the target VPE-Gapping contrast based on proficiency. This analysis revealed that the late L2ers tended to have higher proficiency than the early L2ers; the late L2 group had four participants in the Lower group, 14 in the Medium group, and 12 in the Higher group, whereas the early L2 group had 11 participants in the Lower group, 9 in the Medium group, and 7 in the Higher group. Furthermore, a *t*-test revealed that the late L2ers (as a group) had significantly higher final proficiency scores than early L2ers (as a group) did ($t(55) = 2.548$, $p = .014$, Cohen's $d = 0.690$); the late L2ers also showed significantly higher scores for two of the three components of the proficiency measure, i.e., syntactic complexity ($t(55) = 2.287$,

p = .026, Cohen's d = 0.607) and lexical complexity (t(55) = 2.307, p = .026, Cohen's d = 0.626) (cf. accuracy (t(55) = 1.213, p = .231, Cohen's d = 0.322)).

For the analysis of the RT data, I first removed all raw RTs that were 30 seconds or longer in duration, which led to the removal of 0.24% of the L1 adult data, 0.26% of the L1 child data, 0.00% of the early L2 data, and 0.14% of the late L2 data. Then I identified and removed outliers, which I defined as any RTs deviating from the participant's mean for that condition by more than 1.5 standard deviations. This removal affected 6.28% of the L1 adult data, 5.77% of the L1 child data, 5.93% of the early L2 data, and 4.89% of the late L2 data. Similarly, outliers identified per item based on the mean for each condition were removed, which affected 8.04% of the L1 adult data, 9.61% of the L1 child data, 7.36% of the early L2 data, and 7.78% of the late L2 data. After this data trimming procedure, a linear mixed-effects model was constructed on the RTs. This model included *Construction* and *Clause* as binary fixed effects (contrast-coded with [−.5, .5]) and *participant* and *item* as random effects. For the L1 child data and both sets of the L2 data, additional linear mixed-effects models were constructed for each group to test for potential effects of *Age* (for the L1 child data) and *Proficiency* (for the L2 data). These models included *Construction* and *Clause* as binary fixed effects, *Age* (for the L1 child data) and *Proficiency* (for the L2 data) as continuous fixed effects, and *participant* and *item* as random effects.

All regression models in this dissertation were constructed with the maximal random effects structure allowed by the design (Barr, Levy, Scheepers, & Tily, 2013; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017; Stroup, 2012). The model formula for each analysis can be found in the results table. The modeling was implemented in R (R Core Team, 2018).

5.3 Results

Table 5.2 and Figure 5.2 display the means (and standard deviations) of the acceptance rates for the four conditions by the four participant groups. As predicted, the L1 adults accepted the VPE-C condition, the VPE-A condition, and the Gapping-C condition but rejected the Gapping-A condition. However, the other three groups showed a different judgment pattern; while they accepted VPE regardless of clause type, just as the L1 adults did, their acceptance rates for the Gapping-C condition were not as high as that of the L1 adults. Furthermore, the L1 children and the early L2ers did not fully reject the Gapping-A condition. This pattern of results

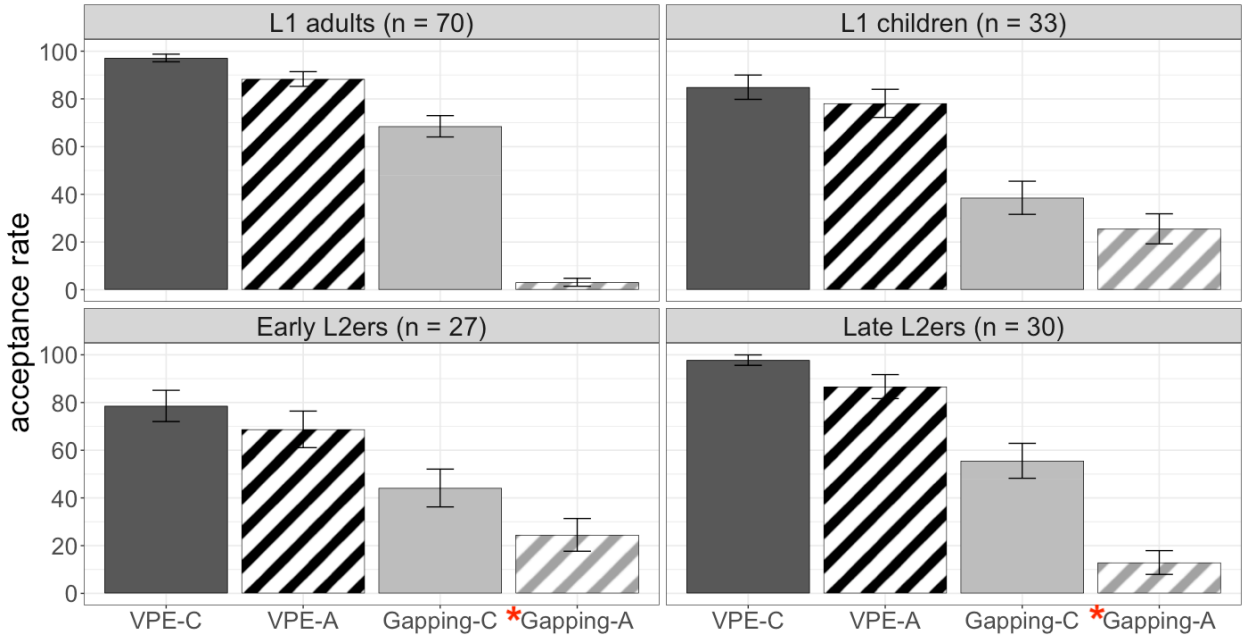
suggests that (on average) the three learner groups did not know that Gapping is grammatical in a conjunct clause and that (on average) the L1 children and early L2ers did not know that Gapping is ungrammatical in an adjunct clause (although it appears that as groups, they did distinguish between the Gapping-C and Gapping-A conditions in the right direction).

Table 5.2
Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Group

	VPE-C	**VPE-A**	**Gapping-C**	***Gapping-A**
L1 Adults	97.14	88.34	68.50	3.11
(*n* = 70)	(*SD* = 16.68)	(*SD* = 32.14)	(*SD* = 46.51)	(*SD* = 17.38)
L1 Children	84.90	78.12	38.54	25.53
(*n* = 33)	(*SD* = 35.90)	(*SD* = 41.45)	(*SD* = 48.80)	(*SD* = 43.72)
Early L2ers	78.57	68.75	44.16	24.52
(*n* = 27)	(*SD* = 41.17)	(*SD* = 46.51)	(*SD* = 49.82)	(*SD* = 43.16)
Late L2ers	97.77	86.67	55.56	12.92
(*n* = 30)	(*SD* = 14.82)	(*SD* = 34.09)	(*SD* = 49.83)	(*SD* = 33.64)

Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause.

Figure 5.2
Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Group



Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause. Error bars represent 95% confidence intervals.

Concerning the asymmetries detected between L1 adults and the other participant groups, a question arises as to why the latter failed to show the target response pattern. One possible and simple answer is that neither the L1 children nor the L2ers have target knowledge of the grammaticality contrast. Alternatively, it is possible that some of them, such as the older L1 children and the L2ers with higher proficiency, had target-like knowledge that was obscured by group averaging. By-group statistical analyses were conducted to test for possible *Age* effects among the L1 children and for *Proficiency* effects among the L2ers as well as to test for differences among the four conditions for each group. In the following section, I begin by reporting the results of the analysis on the L1 adults' responses on the AJT.

5.3.1 L1 adults.

The results of the logistic mixed-effects regression model fitted to the L1 adults' judgment data are summarized in Table 5.3.

Table 5.3
Results of the Logistic Mixed-Effects Regression for the L1 Adult Judgment Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.265	0.731	.084
Construction	−8.751	1.303	< .001
Clause	−7.069	1.311	< .001
Construction × Clause	−5.343	2.609	.041

Note. Model formula: glmer(Judgment ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 | item))

The model found a significant effect of *Construction* ($\beta = -8.751$, $SE = 1.303$, $p < .001$), with higher acceptance in the VPE conditions than in the Gapping conditions, and a significant effect of *Clause* ($\beta = -7.069$, $SE = 1.311$, $p < .001$), with higher acceptance in the Conjunct conditions than in the Adjunct conditions. Importantly, there was a significant interaction between *Construction* and *Clause* ($\beta = -5.343$, $SE = 2.609$, $p = .041$). To identify the source of the interaction, planned pairwise comparisons were performed using a separate mixed-effects model. These analyses found that the L1 adults accepted the VPE-C condition significantly more often than both the VPE-A condition ($\beta = -22.755$, $SE = 6.882$, $p < .001$) and the Gapping-C condition ($\beta = -22.241$, $SE = 6.943$, $p = .001$). This is an interesting result given that all three of these conditions are grammatical and were accepted at far higher rates than the ungrammatical

Gapping-A condition (see Figure 5.2). One possible explanation of this result is that the VPE-A condition and the Gapping-C condition are more taxing than the VPE-C condition in terms of processing. This issue is discussed further in §5.4.1. Critically, the L1 adults accepted the Gapping-A condition significantly less often than both the VPE-A condition ($\beta = -13.560$, $SE = 2.431$, $p < .001$) and the Gapping-C condition ($\beta = -10.208$, $SE = 1.766$, $p < .001$); furthermore, the acceptance rate in the Gapping-C condition was close to "0" (i.e., 3.11%). In sum, these results indicate that L1 adults have the target contrast between VPE and Gapping, as predicted.

5.3.2 L1 children.

Table 5.4 presents the results of the mixed-effects regression analysis on the L1 children's judgment data.

Table 5.4
Results of the Logistic Mixed-Effects Regression for the L1 Child Judgment Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.526	0.245	.032
Construction	−3.209	0.454	< .001
Clause	−0.912	0.351	.009
Construction × Clause	−0.102	0.606	.866

Note. Model formula: glmer(Judgment ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

This analysis revealed a significant effect of *Construction* ($\beta = -3.209$, $SE = 0.454$, $p = < .001$) and a significant effect of *Clause* ($\beta = -0.912$, $SE = 0.351$, $p = .009$), with higher acceptance for the VPE conditions than the Gapping conditions and higher acceptance for the Conjunct conditions than the Adjunct conditions. However, there was no significant interaction between *Construction* and *Clause* ($\beta = -0.102$, $SE = 0.606$, $p = .866$).

To test the possibility that the large standard deviations obscured the fact that the L1 children treated the four conditions differently (see Table 5.2), I conducted planned pairwise comparisons using mixed-effects models for each *Construction* and *Clause*. These analyses revealed that L1 children accepted the VPE-C condition at higher rates than both the VPE-A condition ($\beta = -0.809$, $SE = 0.465$, $p = .082$) and the Gapping-C condition ($\beta = -3.019$, $SE = 0.576$, $p < .001$); however, the difference between the VPE-C condition and the VPE-A

condition was only marginally significant. Furthermore, L1 children accepted the Gapping-A condition significantly less often than both the VPE-A condition ($\beta = -3.158$, $SE = 0.465$, $p < .001$) and the Gapping-C condition ($\beta = -0.932$, $SE = 0.398$, $p = .019$). These results show that the L1 children as a group were able to reject Gapping in adjunct clauses.

A potential effect of *Age* was investigated by constructing a separate logistic mixed-effects regression model with the fixed effect *Age* added. This model did not reveal a main effect of *Age* ($\beta = 0.159$, $SE = 0.183$, $p = .386$), as shown in Table 5.5. Nor was there a statistically significant three-way interaction among *Construction*, *Clause*, and *Age* ($\beta = -0.198$, $SE = 0.442$, $p = .654$).

Table 5.5
Results of the Logistic Mixed-Effects Regression for the L1 Child Judgment Data in Study 2 with the Factor Age *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	−0.481	1.150	.676
Construction	4.861	1.898	.010
Clause	2.652	1.550	.087
Age	0.159	0.183	.386
Construction × Clause	1.148	2.643	.664
Construction × Age	−1.271	0.307	< .001
Clause × Age	−0.558	0.255	.029
Construction × Clause × Age	−0.198	0.442	.654

Note. Model formula: glmer(Judgment ~ Construction * Clause * Age + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

However, the factor *Age* did significantly interact with both *Construction* ($\beta = -1.271$, $SE = 0.307$, $p < .001$) and *Clause* ($\beta = -0.558$, $SE = 0.255$, $p = .029$). Follow-up simple regression analyses revealed that older L1 children displayed higher acceptance for the VPE conditions over the Gapping conditions ($\beta = 36.494$, $SE = 9.086$, $p < .001$) as well as for the Conjunct conditions over the Adjunct conditions ($\beta = 14.498$, $SE = 7.136$, $p = .051$).

The significant interactions found between *Age* and *Construction* and between *Age* and *Clause* and the large standard deviations in the acceptance rates for the Gapping-C condition ($M = 38.54$; $SD = 48.80$) all led me to wonder whether any effect of *Age* would emerge under a different type of analysis. To address this question, I ran a simple regression analysis on the data from the L1 children with *Age* as an independent variable and the strength of the sensitivity to

the target grammaticality contrast as a dependent variable. I measured this sensitivity for each participant by subtracting the mean acceptance rate of the ungrammatical condition (Gapping-A) from the average of the mean acceptance rates for all grammatical conditions (VPE-C, VPE-A, Gapping-C), using the equation in (7). The sensitivity scores for all the participants are provided in Appendix F.

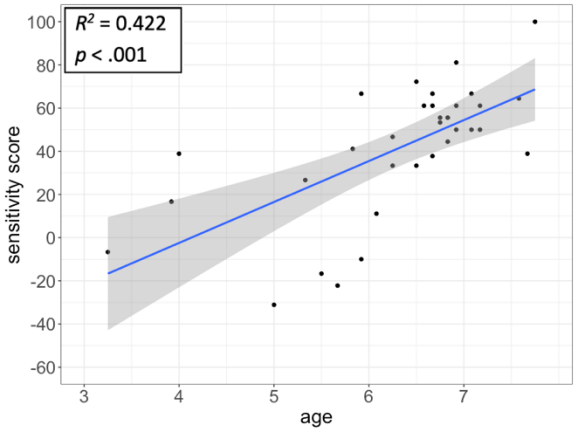
(7) Grammaticality contrast sensitivity scores

([Mean acceptance rate for VPE-C] + [Mean acceptance rate for VPE-A]
+ [Mean acceptance rate for Gapping-C]) / 3 − [Mean acceptance rate for *Gapping-A]

A sensitivity score can range from −100 to 100. Higher scores indicate stronger sensitivity to the target grammaticality contrast. A score of 0 indicates that the participant treats the Gapping-A condition as equal to the other conditions in terms of acceptability.

As shown in Figure 5.3, the results of the simple regression on the sensitivity scores showed a significant effect of *Age* ($\beta = 18.964$, $SE = 3.988$, $p < .001$), thus indicating that older children have higher sensitivity to the target grammaticality contrast between VPE and Gapping.

Figure 5.3
Relation between Age and Sensitivity Score for the L1 Children in Study 2



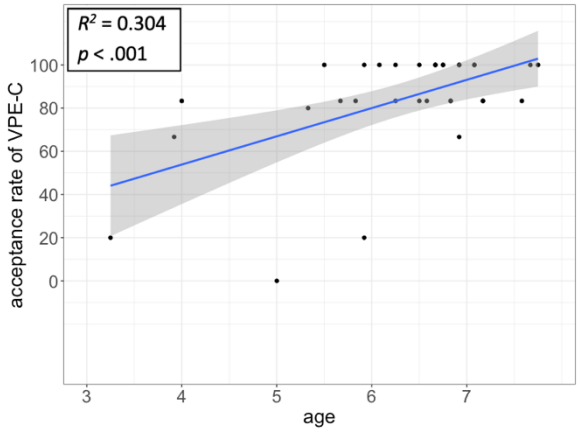
Note. The shaded region shows the 95% confidence interval for the sample mean.

Further simple regression analyses revealed that this effect of *Age* on the sensitivity scores mainly came from older children's significantly higher acceptance rates for the VPE-C condition ($\beta = 13.073$, $SE = 3.552$, $p < .001$) and the VPE-A condition ($\beta = 8.774$, $SE = 4.030$, $p = .037$)

and their significantly lower acceptance rate for the Gapping-A condition ($\beta = -12.422$, $SE = 4.046$, $p = .004$), as exhibited in Figures 5.4–5.6.

Figure 5.4

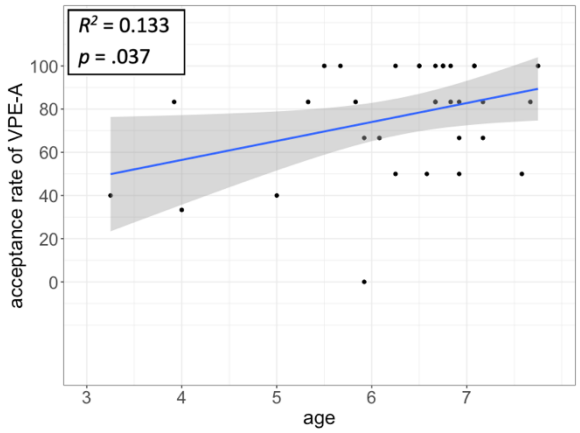
Relation between Age and Acceptance Rate of VPE-C for the L1 Children in Study 2



Note. The shaded region shows the 95% confidence interval for the sample mean.

Figure 5.5

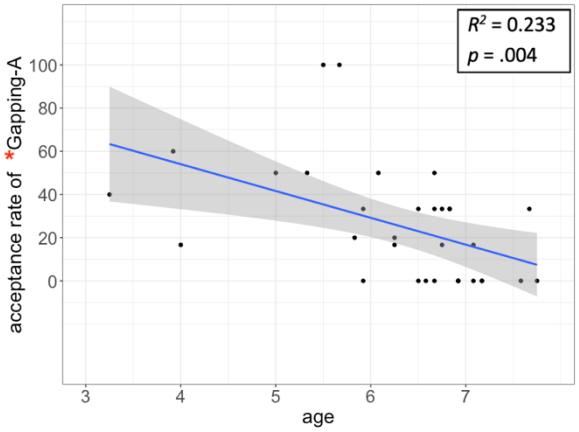
Relation between Age and Acceptance Rate of VPE-A for the L1 Children in Study 2



Note. The shaded region shows the 95% confidence interval for the sample mean.

Figure 5.6

*Relation between Age and Acceptance Rate of *Gapping-A for the L1 Children in Study 2*



Note. The shaded region shows the 95% confidence interval for the sample mean.

However, the analyses found no significant effect of *Age* on the acceptance rates for the Gapping-C condition ($\beta = -2.223$, $SE = 5.064$, $p = .664$), thus indicating that the L1 children accepted this condition to more or less the same degree regardless of *Age*. Overall, older L1 children were shown to possess the target grammaticality contrast between VPE and Gapping.

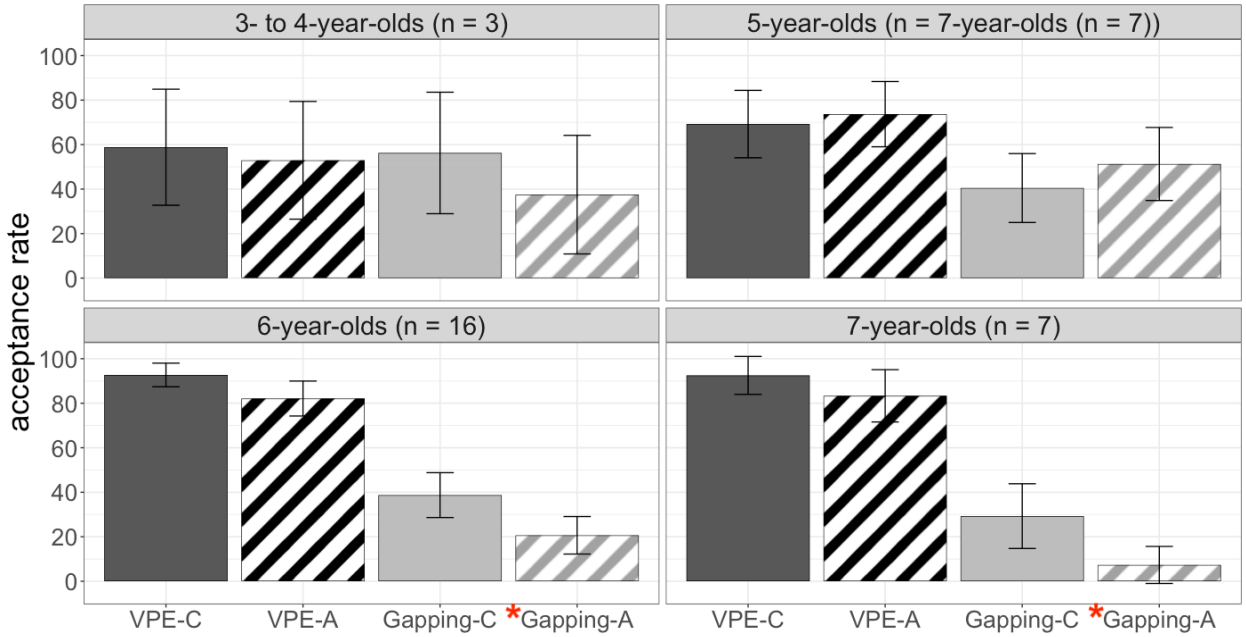
To pinpoint the age at which the L1 children started to display the VPE-Gapping contrast, I divided them into four groups: 3 to 4-year-olds ($n = 3$), 5-year-olds ($n = 7$), 6-year-olds ($n = 16$), and 7-year-olds ($n = 7$). The 3-year-olds ($n = 2$) and the 4-year-old ($n = 1$) were collapsed into a single group because there were so few, and none showed the target contrast. I then analyzed the judgment patterns of each age group. Table 5.6 and Figure 5.7 show the results of this analysis.

Table 5.6
L1 Children's Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Age Group

	VPE-C	**VPE-A**	**Gapping-C**	***Gapping-A**
3- to 4-year-olds	58.82	52.94	56.25	37.50
(*n* = 3)	(*SD* = 50.73)	(*SD* = 51.45)	(*SD* = 51.23)	(*SD* = 50.00)
5-year-olds	69.23	73.68	40.48	51.28
(*n* = 7)	(*SD* = 46.76)	(*SD* = 44.63)	(*SD* = 49.68)	(*SD* = 50.64)
6-year-olds	92.71	82.11	38.71	20.65
(*n* = 16)	(*SD* = 26.14)	(*SD* = 38.53)	(*SD* = 48.97)	(*SD* = 40.70)
7-year-olds	92.50	83.33	29.27	7.32
(*n* = 7)	(*SD* = 26.67)	(*SD* = 37.72)	(*SD* = 46.06)	(*SD* = 26.37)

Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause.

Figure 5.7
L1 Children's Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Age Group



Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause. Error bars represent 95% confidence intervals.

There was no target-like pattern observed for the group analyses of the 3- to 4-year-olds or the 5-year-olds. By contrast, the 6-year-olds and the 7-year-olds treated the Gapping-A condition differently from the other conditions and accepted the VPE conditions. In particular, the latter group clearly rejected the Gapping-A condition. However, the 6- to 7-year-old L1 children

accepted the Gapping-C condition only about 30%–40% of the time. I will argue in §5.4.1 and §5.4.2 that this pattern of results stems from processing difficulty.

Overall, the L1 children appeared to have acquired the target contrast by the age of 7, at least when the results are analyzed by age group. However, I will show in §5.4.2 that a close inspection of the individual data reveals that some of the younger L1 children also evince the target grammaticality contrast.

5.3.3 Early L2ers.

The logistic mixed-effects model for the early L2ers produced the results in Table 5.7.

Table 5.7
Results of the Logistic Mixed-Effects Regression for the Early L2er Judgment Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.110	0.254	.664
Construction	−2.468	0.525	< .001
Clause	−0.969	0.338	.004
Construction × Clause	−0.749	0.595	.208

Note. Model formula: glmer(Judgment ~ Construction * Clause + (1 + Construction * Clause | participant) + (1+ Construction * Clause | item))

The model showed a significant effect of both *Construction* ($\beta = -2.468$, $SE = 0.525$, $p < .001$) and *Clause* ($\beta = -0.969$, $SE = 0.338$, $p = .004$), with higher acceptance rates for the VPE conditions than for the Gapping conditions and higher acceptance rates for the Conjunct conditions than the Adjunct conditions. However, there was no significant interaction between *Construction* and *Clause* ($\beta = -0.749$, $SE = 0.595$, $p = .208$).

The large standard deviations in the acceptance rates for the Gapping-C condition ($M = 44.16$; $SD = 49.82$) may have been the reason that a significant interaction between *Construction* and *Clause* did not emerge, as was the case for the L1 children. I therefore ran separate mixed-effects models on the data from each *Construction* and *Clause* for pairwise comparisons. This analysis revealed that the Gapping-A condition was accepted significantly less often than both the VPE-A condition ($\beta = -2.663$, $SE = 0.554$, $p < .001$) and the Gapping-C condition ($\beta = -1.384$, $SE = 0.429$, $p = .001$), indicating that the early L2ers as a group were able to reject Gapping in adjunct clauses. However, I will show below and in §5.4.2 that only the early L2ers with higher proficiency had knowledge of the ungrammaticality of Gapping in

adjunct clauses. In addition, the early L2ers accepted the Gapping-C condition significantly less often than the VPE-C condition ($\beta = -2.049$, $SE = 0.541$, $p < .001$), presumably due to the processing difficulty associated with the Gapping-C condition (see §5.4.1). Their acceptance rates did not differ between the VPE-C condition and the VPE-A condition ($\beta = -0.536$, $SE = 0.430$, $p = .212$).

An additional logistic mixed-effects regression analysis was conducted to test if there was any effect of *Proficiency* on the early L2ers' judgment patterns. The model with the *Proficiency* factor added did not reveal a significant effect of *Proficiency* ($\beta = -0.078$, $SE = 0.082$, $p = .341$), as shown in Table 5.8.

Table 5.8
Results of the Logistic Mixed-Effects Regression for the Early L2er Judgment Data in Study 2 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.072	0.262	.784
Construction	−2.924	0.457	< .001
Clause	−0.942	0.370	.011
Proficiency	−0.078	0.082	.341
Construction × Clause	−0.735	0.652	.260
Construction × Proficiency	−0.609	0.155	< .001
Clause × Proficiency	−0.024	0.113	.834
Construction × Clause × Proficiency	−0.049	0.204	.809

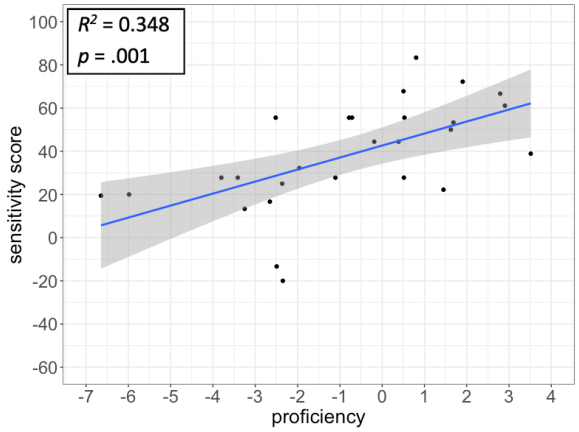
Note. Model formula: glmer(Judgment ~ Construction * Clause * Proficiency + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

The interaction between *Clause* and *Proficiency* ($\beta = -0.024$, $SE = 0.113$, $p = .834$) and the three-way interaction among *Construction*, *Clause*, and *Proficiency* ($\beta = -0.049$, $SE = 0.204$, $p = .809$) did not reach statistical significance, either. However, a significant interaction between *Construction* and *Proficiency* emerged ($\beta = -0.609$, $SE = 0.155$ $p < .001$). A further simple regression analysis found that the factor *Proficiency* had a significant effect on the degree to which the early L2ers discriminated between the VPE and Gapping conditions, with higher-proficiency participants showing higher acceptance for VPE than Gapping ($\beta = 18.541$, $SE = 4.095$, $p < .001$).

Next, to further investigate the effect of *Proficiency*, I ran a simple regression analysis on the sensitivity scores (computed using the equation in (7); to view the sensitivity scores for each

of the early L2ers, see Appendix F) with *Proficiency* as an independent variable. This analysis revealed a significant effect of *Proficiency* (β = 5.559, *SE* = 1.521, p = .001), trending toward higher sensitivity to the target contrast in higher-proficiency early L2ers (see Figure 5.8).

Figure 5.8
Relation between Proficiency and Sensitivity Score for the Early L2ers in Study 2

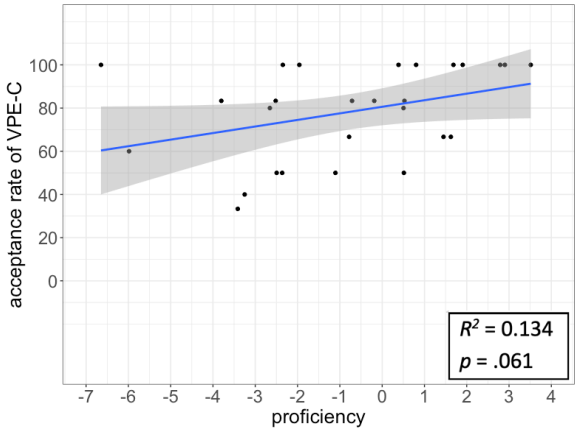


Note. The shaded region shows the 95% confidence interval for the sample mean.

Crucially, further simple regression analyses for each of the four conditions showed that *Proficiency* was a marginally significant predictor of acceptance rates for both the VPE-C condition (β = 3.041, *SE* = 1.547, p = .061) and the VPE-A condition (β = 3.781, *SE* = 2.037, p = .075). As shown in Figures 5.9 and 5.10, early L2ers with higher proficiency showed a significantly higher acceptance rate for, respectively, the VPE-C condition and the VPE-A condition.

Figure 5.9

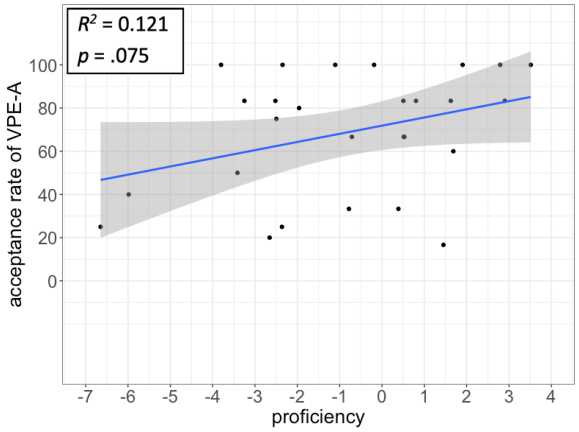
Relation between Proficiency and Acceptance Rate of VPE-C for the Early L2ers in Study 2



Note. The shaded region shows the 95% confidence interval for the sample mean.

Figure 5.10

Relation between Proficiency and Acceptance Rate of VPE-A for the Early L2ers in Study 2

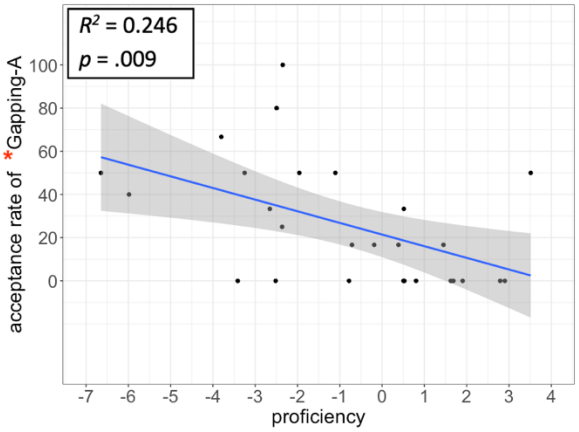


Note. The shaded region shows the 95% confidence interval for the sample mean.

Importantly, early L2ers with higher proficiency showed a significantly lower acceptance rate for the ungrammatical Gapping-A condition ($\beta = -5.393$, $SE = 1.889$, $p = .009$), as shown in Figure 5.11.

Figure 5.11

*Relation between Proficiency and Acceptance Rate of *Gapping-A for the Early L2ers in Study 2*

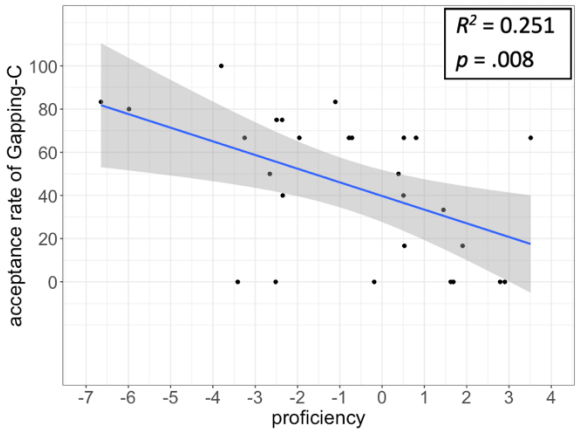


Note. The shaded region shows the 95% confidence interval for the sample mean.

However, *Proficiency* was also found to be a significant predictor of acceptance rates for the grammatical Gapping-C condition ($\beta = -6.326$, $SE = 2.184$, $p = .008$) such that higher-proficiency L2ers again showed less acceptance in this condition (see Figure 5.12).

Figure 5.12

Relation between Proficiency and Acceptance Rate of Gapping-C for the Early L2ers in Study 2



Note. The shaded region shows the 95% confidence interval for the sample mean.

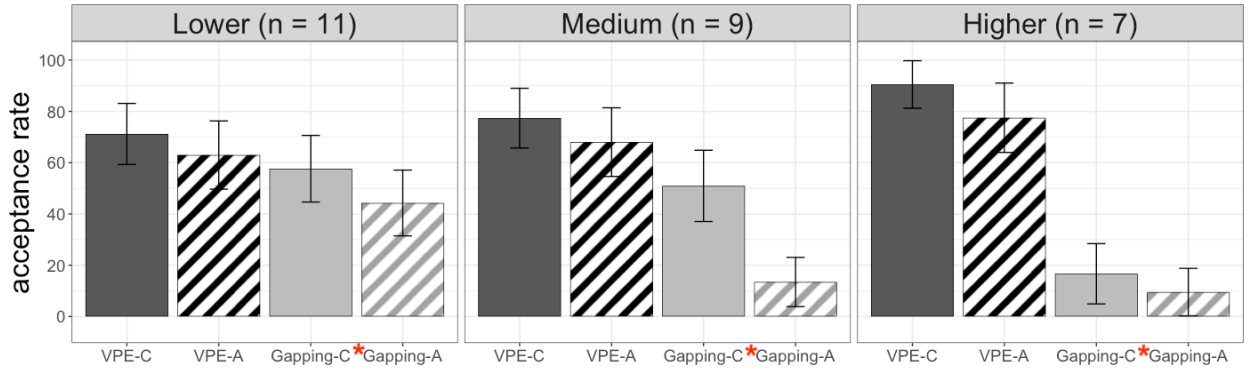
To try to capture a clearer picture of proficiency effects in the data from the early L2ers, I grouped them into three proficiency groups based on their proficiency scores (see Appendix A) and examined how each group performed on the AJT. The results are summarized in Table 5.9 and Figure 5.13.

Table 5.9
Early L2ers' Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Proficiency Group

	VPE-C	**VPE-A**	**Gapping-C**	***Gapping-A**
Lower	71.19	62.96	57.63	44.26
(*n* = 11)	(*SD* = 45.68)	(*SD* = 48.74)	(*SD* = 49.84)	(*SD* = 50.08)
Medium	77.36	68.00	50.94	13.46
(*n* = 9)	(*SD* = 42.25)	(*SD* = 47.12)	(*SD* = 50.47)	(*SD* = 34.46)
Higher	90.48	77.50	16.67	9.52
(*n* = 7)	(*SD* = 29.71)	(*SD* = 42.29)	(*SD* = 37.72)	(*SD* = 29.71)

Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause.

Figure 5.13
Early L2ers' Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Proficiency Group



Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause. Error bars represent 95% confidence intervals.

While the Lower group showed a weak contrast between VPE and Gapping, this group did not seem to reject the Gapping-A condition fully. By contrast, the Medium and Higher groups successfully rejected the Gapping-A condition and accepted the VPE-C and VPE-A conditions. However, the Higher group's acceptance rate for the Gapping-C condition is very low. In §5.4.1 and §5.4.2, I will argue that this pattern of results does not necessarily indicate that the participants lacked knowledge of the grammaticality of Gapping in a conjunct clause.

Taken together, the results seem to indicate that early L2ers with higher proficiency are able to acquire the target-like contrast between VPE and Gapping.

5.3.4 Late L2ers.

Table 5.10 summarizes a logistic mixed-effects model run for the late L2ers' data.

Table 5.10
Results of the Logistic Mixed-Effects Regression for the Late L2er Judgment Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.654	1.163	.155
Construction	−7.642	2.187	< .001
Clause	−4.077	2.910	.161
Construction × Clause	−3.100	5.709	.587

Note. Model formula: glmer(Judgment ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

The model showed a significant effect of *Construction* ($\beta = -7.642$, $SE = 2.187$, $p < .001$), with higher acceptance rates for the VPE conditions than for the Gapping conditions. There was no significant effect of *Clause* ($\beta = -4.077$, $SE = 2.910$, $p = .161$). Even though analyses with descriptive statistics clearly showed that the late L2ers possessed the target contrast (see Table 5.2 and Figure 5.2), the mixed-effects model found no interaction between *Construction* and *Clause* ($\beta = -3.100$, $SE = 5.709$, $p = .587$). I argue that this null result, like those for the L1 child group and the early L2 group, resulted from the large standard deviations for the acceptance rates in the Gapping-C condition ($M = 55.56$; $SD = 49.83$). In fact, separate mixed-effects models for pairwise comparisons revealed that late L2ers accepted the Gapping-A condition significantly less often than both the VPE-A condition ($\beta = -10.252$, $SE = 2.791$, $p < .001$) and the Gapping-C condition ($\beta = -5.439$, $SE = 1.172$, $p < .001$). This result, together with the low acceptance rate of the Gapping-A condition in the late L2ers ($M = 12.92$; $SD = 33.64$), indicates that they know that Gapping in adjunct clauses is ungrammatical. Their judgments did not show a significant difference between the VPE-C condition and the VPE-A condition ($\beta = -5.006$, $SE = 3.745$, $p = .181$) or between the VPE-C condition and the Gapping-C condition ($\beta = -9.475$, $SE = 8.627$, $p = .272$).

To identify any possible effect of *Proficiency*, I ran a separate logistic mixed-effects model with *Proficiency* added as a continuous fixed effect. As shown in Table 5.11, this model found no significant effect of *Proficiency* ($\beta = -0.333$, $SE = 0.277$, $p = .229$). Nor was there a significant interaction between *Construction* and *Proficiency* ($\beta = -0.724$, $SE = 0.477$, $p = .129$),

between *Clause* and *Proficiency* ($\beta = 0.535$, $SE = 0.383$, $p = .162$), or among *Construction*, *Clause*, and *Proficiency* ($\beta = -0.150$, $SE = 0.728$, $p = .837$).

Table 5.11
Results of the Logistic Mixed-Effects Regression for the Late L2er Judgment Data in Study 2 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.612	1.009	.110
Construction	−6.505	1.815	< .001
Clause	−4.125	2.170	.057
Proficiency	−0.333	0.277	.229
Construction × Clause	−3.267	4.293	.447
Construction × Proficiency	−0.724	0.477	.129
Clause × Proficiency	0.535	0.383	.162
Construction × Clause × Proficiency	−0.150	0.728	.837

Note. Model formula: glmer(Judgment ~ Construction * Clause * Proficiency + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

The effect of *Proficiency* in the data was further investigated using a different analysis method. I conducted a simple regression analysis on the sensitivity scores (computed using the equation in (7); for the individual sensitivity scores, see Appendix F) with *Proficiency* as the independent variable. This analysis did not reveal any significant effect of *Proficiency* on the sensitivity scores ($\beta = 0.580$, $SE = 2.233$, $p = .797$), the acceptance rate for the VPE-C condition ($\beta = -0.547$, $SE = 0.770$, $p = .483$), the acceptance rate for the VPE-A condition ($\beta = 3.761$, $SE = 2.210$, $p = .100$), or the acceptance rate for the Gapping-A condition ($\beta = -2.188$, $SE = 2.530$, $p = .395$). However, a marginally significant effect of *Proficiency* was found for the Gapping-C condition ($\beta = -8.037$, $SE = 4.201$, $p = .066$), trending toward less acceptance as proficiency increases. This finding for the grammatical Gapping-C condition will be discussed in terms of processing in §5.4.1 and §5.4.2.

Lastly, as with the early L2ers' data, I examined the late L2ers' data by dividing them on the basis of their proficiency scores (see Appendix A). Table 5.12 and Figure 5.14 show that while all the three proficiency groups showed the grammaticality contrast between VPE and Gapping, the Lower group did not fully reject the ungrammatical Gapping-A condition. As the late L2ers' proficiency scores went up, their acceptance rates for both the (grammatical)

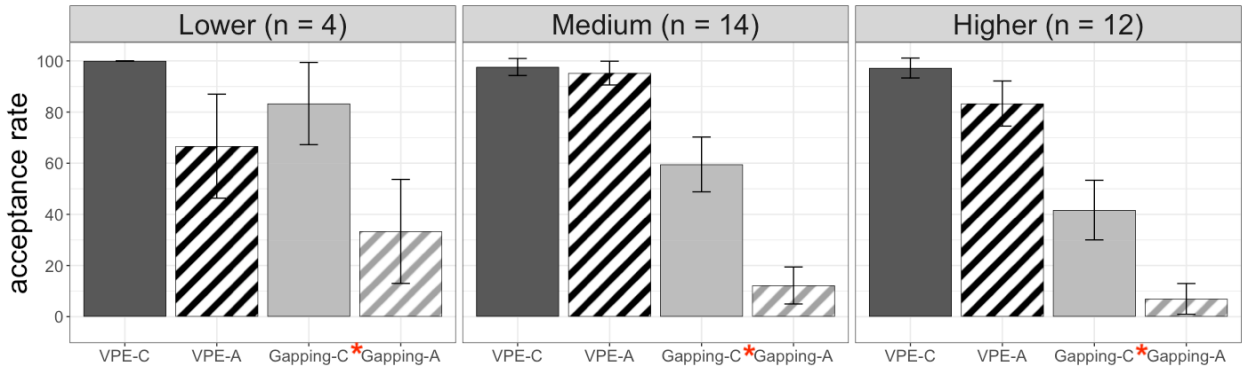
Gapping-C condition and the (ungrammatical) Gapping-A condition tended to decrease. This result is in line with the data from the early L2ers.

Table 5.12
Late L2ers' Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Proficiency Group

	VPE-C	**VPE-A**	**Gapping-C**	***Gapping-A**
Lower	100.00	66.67	83.33	33.33
(*n* = 4)	(*SD* = 0.00)	(*SD* = 48.15)	(*SD* = 38.07)	(*SD* = 48.15)
Medium	97.62	95.24	59.52	12.20
(*n* = 14)	(*SD* = 15.34)	(*SD* = 21.42)	(*SD* = 49.38)	(*SD* = 32.92)
Higher	97.22	83.33	41.67	6.94
(*n* = 12)	(*SD* = 16.55)	(*SD* = 37.53)	(*SD* = 49.65)	(*SD* = 25.60)

Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause.

Figure 5.14
Late L2ers' Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Proficiency Group



Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause. Error bars represent 95% confidence intervals.

In sum, the late L2ers in the Medium and Higher groups seem to have the target grammaticality contrast between VPE and Gapping; they displayed a strong tendency to accept VPE in conjunct and adjunct clauses and to reject Gapping in adjunct clauses.

5.4. Discussion

This study investigated the acquisition of the grammaticality contrast between VPE and Gapping by L1-English children and L1-Korean L2ers of English. The results from the AJT

showed that the older L1 children, the higher-proficiency early L2ers, and the late L2ers had succeeded in acquiring the target contrast. I begin this section by examining why the L1 adult controls did not accept the VPE-A and Gapping-C conditions at the same rates that they accepted the VPE-C condition. Based on my analysis of the RTs and the response patterns of individual participants, I argue that participants experienced processing difficulty when encountering VPE in adjunct clauses and Gapping in conjunct clauses and that the latter was the harder of the two to process. Then I move on to propose possible L1 and L2 developmental sequences based on my analysis of the data from individual participants, using *Age* and *Proficiency* as guidelines for, respectively, the L1 children and the L2ers. Lastly, I discuss limitations and directions of future research.

5.4.1 Processing difficulty of VP-Ellipsis in adjunct clauses and Gapping in conjunct clauses.

L1 adults and L1 children accepted both the VPE-C condition and VPE-A condition but accepted the latter significantly less often than the former. I propose that the presence of the logical subordinator *because* may have contributed to the low acceptance rates in the VPE-A condition. This subordinator is associated with *cause* and *effect* relationships, inviting complex causal reasoning (Evers-Vermeul & Sanders, 2009; Kuiken & Vedder, 2012; Spooren & Sanders, 2008). It is thus conceivable that the lower acceptance observed with the VPE-A sentences (e.g., *Sara made pizza because Kelly did*) stemmed from an added cognitive burden imposed by the need to deal with complex semantic relationships.

The Gapping-C condition was also accepted less often than the VPE-C condition by L1 adults, L1 children, and early L2ers. I propose that this pattern of results can be explained in terms of syntactic processing. For one thing, Gapping is a structurally complex phenomenon involving Across-the-Board movement (see §2.2). The parsing of Gapping sentences thus requires language users to construct a dependency between a moved verb filler and its gap, which is likely to be a taxing process. The fact that such verb fillers cannot serve as a cue for movement, in contrast to the case of moved *wh*-fillers in questions and relative clauses, could make it even harder for language users to resolve the filler–gap dependency. Furthermore, the discourse contexts in which Gapping is felicitous are generally more restricted than those in which VPE is felicitous; whereas Gapping sounds rather odd when uttered out of the blue, VPE

can be felicitous even when it appears at the beginning of a discourse.[4] In addition, the absence of an overt verb in Gapping may be highly noticeable (especially in written format).

To investigate whether the above processing explanation should be adopted, I examined the RT data for each participant group. This analysis assumed that longer RTs indicate greater processing difficulty stemming from involved propositional relations, complex syntactic structures, encoding/retrieval interferences, etc., as the existing processing literature has suggested (e.g., for an eye-tracking-while-reading paradigm, see Kliegl & Laubrock, 2018; for a grammaticality judgment paradigm, see Smith, 2011; for a self-paced listening picture-verification paradigm, see Peristeri & Tsimpli, 2013; for a self-paced reading paradigm, see Villata, Tabor, & Franck, 2018). Table 5.13 and Figure 5.15 show that all the learner groups judged the VPE-C condition faster than all other conditions and that the L1 adults judged the VPE-C condition faster than the VPE-A and Gapping-C conditions; together, these data suggest that VPE in conjunct clauses was relatively easy to process.

Table 5.13
Mean Reaction Times (in Seconds) in the Acceptability Judgment Task per Condition and Group

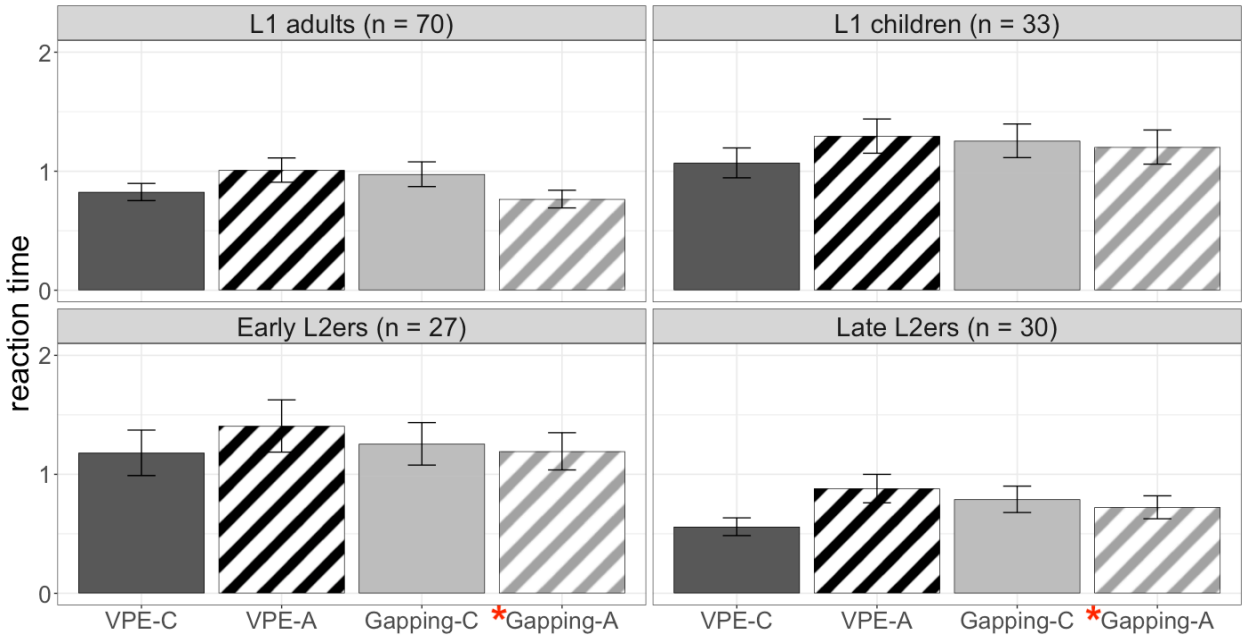
	VPE-C	**VPE-A**	**Gapping-C**	***Gapping-A**
L1 Adults	0.83	1.01	0.97	0.77
(*n* = 70)	(*SD* = 0.72)	(*SD* = 0.97)	(*SD* = 0.98)	(*SD* = 0.73)
L1 Children	1.07	1.30	1.26	1.20
(*n* = 33)	(*SD* = 0.83)	(*SD* = 0.94)	(*SD* = 0.90)	(*SD* = 0.91)
Early L2ers	1.18	1.41	1.26	1.19
(*n* = 27)	(*SD* = 1.11)	(*SD* = 1.23)	(*SD* = 1.04)	(*SD* = 0.93)
Late L2ers	0.56	0.88	0.79	0.72
(*n* = 30)	(*SD* = 0.49)	(*SD* = 0.73)	(*SD* = 0.71)	(*SD* = 0.61)

Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause.

[4] I thank Bonnie D. Schwartz (personal communication, 4 April 2020) for bringing this to my attention.

Figure 5.15

Mean Reaction Times (in seconds) in the Acceptability Judgment Task per Condition and Group



Notes. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause. Error bars represent 95% confidence intervals.

For the L1 adults, it was the ungrammatical Gapping-A condition that displayed the shortest RTs, thereby indicating that they did not experience processing difficulty when judging these sentences. One might argue that this result is surprising because self-paced reading studies and eye-tracking-while-reading studies have generally found increased RTs at the segment that language users find ungrammatical (e.g., Biondo, Vespignani, & Dillon, 2019) or implausible (e.g., Omaki, Lau, White, Dakan, Apple, & Phillips, 2015). However, my task was a grammaticality judgment task in which the target sentences were orally presented twice (with their written counterparts), and it therefore imposed less time pressure on participants than standard self-paced reading or eye-tracking-while-reading tasks. It makes sense that L1 adults, who are presumably more efficient than L1 children (e.g., Wulfeck, 1993) and L2ers (e.g., Hopp, 2018) at integrating linguistic, pragmatic, real-world (etc.) information, would be very quick at detecting when a sentence is ungrammatical and deciding that they need not continue constructing a representation for such strings. In fact, this result is consistent with Smith's (2011) study on passives, which found shorter RTs for ungrammatical sentences (e.g., **The truck was driven the woman.*) than for grammatical ones (e.g., *The truck was driven by the woman.*).

The VPE-A and Gapping-C conditions yielded longer RTs than the other two conditions for the L1 adults and all three learner groups. This indicates that VPE in adjuncts and Gapping in conjuncts were harder to process, as discussed above.

To statistically test for processing effects associated with the factors *Construction* and *Clause*, I constructed a linear mixed-effects model on the RTs for each group. In this section, effects which did not reach (at least marginal) statistical significance will not be discussed, but they are still reported in the results table.

The analysis of the L1 adult data showed a significant interaction between *Construction* and *Clause* ($\beta = -0.453$, $SE = 0.135$, $p = .003$), as shown in Table 5.14.

Table 5.14
Results of the Linear Mixed-Effects Regression for the L1 Adult Reaction Time Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.927	0.053	< .001
Construction	−0.046	0.065	.483
Clause	0.001	0.066	.991
Construction × Clause	−0.453	0.135	.003

Note. Model formula: lmer(RT ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

A follow-up pairwise comparison revealed that RTs were significantly longer in the VPE-A condition than in both the VPE-C condition ($\beta = 0.202$, $SE = 0.058$, $p < .001$) and the Gapping-A condition ($\beta = -0.263$, $SE = 0.069$, $p < .001$). The Gapping-C condition also had significantly longer RTs than both the VPE-C condition ($\beta = 0.181$, $SE = 0.065$, $p = .007$) and the Gapping-A condition ($\beta = -0.211$, $SE = 0.059$, $p < .001$). These results support a processing explanation in which the VPE-A and Gapping-C sentences are harder to process and hence judge than the sentences in the other two conditions. An additional multiple comparison analysis did not reveal a significant difference between the RTs for the VPE-A and Gapping-C conditions ($\beta = -0.046$, $SE = 0.118$, $p = 0.980$).

In contrast to the linear mixed-effects model constructed for the L1 adults, the one for the L1 children did not reveal a main effect of either *Construction* or *Clause*, or a significant interaction between the two factors, as shown in Table 5.15.

Table 5.15
Results of the Linear Mixed-Effects Regression for the L1 Child Reaction Time Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.226	0.066	< .001
Construction	0.051	0.079	.527
Clause	0.096	0.075	.210
Construction × Clause	−0.283	0.195	.157

Note. Model formula: lmer(RT ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

An additional model constructed to test for the *Age* effect in the L1 children's data did not find any significant effects, either (see Table 5.16).

Table 5.16
Results of the Linear Mixed-Effects Regression for the L1 Child Reaction Time Data in Study 2 with the Factor Age *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.529	0.387	.000
Construction	0.020	0.503	.969
Clause	0.285	0.455	.532
Age	−0.048	0.060	.433
Construction × Clause	0.716	1.061	.503
Construction × Age	0.004	0.078	.957
Clause × Age	-0.030	0.070	.673
Construction × Clause × Age	−0.158	0.163	.340

Note. Model formula: lmer(RT ~ Construction * Clause * Age + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

The linear mixed-effects model for the early L2ers found no significant effect of *Construction* or *Clause*, as shown in Table 5.17. The two factors did not interact, either. This result is in line with the result from the L1 child data.

Table 5.17
Results of the Linear Mixed-Effects Regression for the Early L2er Reaction Time Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.380	0.129	< .001
Construction	−0.106	0.097	.282
Clause	0.127	0.114	.277
Construction × Clause	−0.457	0.286	.122

Note. Model formula: lmer(RT ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

However, a separate model for the early L2ers' data with the factor *Proficiency* added showed a significant effect of *Proficiency* (β = −0.104, *SE* = 0.046, p = .031), as shown in Table 5.18. A follow-up simple regression analysis revealed that RTs were shorter for L2ers with higher proficiency scores (β = −0.102, *SE* = 0.051, p = .058), but this trend was only marginally significant.

Table 5.18
Results of the Linear Mixed-Effects Regression for the Early L2er Reaction Time Data in Study 2 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.307	0.127	.000
Construction	−0.120	0.111	.295
Clause	0.066	0.107	.544
Proficiency	−0.104	0.046	.031
Construction × Clause	−0.366	0.271	.190
Construction × Proficiency	0.004	0.036	.902
Clause × Proficiency	−0.118	0.035	.001
Construction × Clause × Proficiency	0.183	0.090	.053

Note. Model formula: lmer(RT ~ Construction * Clause * Proficiency + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

The model furthermore revealed a significant interaction between *Clause* and *Proficiency* (β = −0.118, *SE* = 0.035, p = .001). A follow-up simple regression analysis found that the source of this interaction was that the higher-proficiency early L2ers had longer RTs in the Conjunct conditions than in the Adjunct conditions (β = 0.220, *SE* = 0.074, p = .007). In addition, there was a marginally significant interaction among *Construction*, *Clause*, and *Proficiency* (β = 0.183, *SE* = 0.090, p = .053). Follow-up simple regression analyses found that this interaction came

from the fact that as their proficiency increased, the early L2ers' RTs decreased in the VPE-A condition ($\beta = -0.205$, $SE = 0.071$, $p = .008$), the Gapping-C condition ($\beta = -0.097$, $SE = 0.055$, $p = .092$), and the Gapping-A condition ($\beta = -0.117$, $SE = 0.045$, $p = .015$), but not in the VPE-C condition ($\beta = 0.007$, $SE = 0.069$, $p = .925$).

The results of the linear mixed-effects model for the late L2ers are presented in Table 5.19; they revealed a significant effect of *Clause* ($\beta = 0.155$, $SE = 0.066$, $p = .028$), with longer RTs in the Adjunct conditions than in the Conjunct conditions.

Table 5.19
Results of the Linear Mixed-Effects Regression for the Late L2er Reaction Time Data in Study 2

	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.755	0.062	< .001
Construction	0.011	0.075	.882
Clause	0.155	0.066	.028
Construction × Clause	−0.384	0.152	.018

Note. Model formula: lmer(RT ~ Construction * Clause + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

A significant interaction also emerged between *Construction* and *Clause* ($\beta = -0.384$, $SE = 0.152$, $p = .018$). To unpack the interaction, I ran separate mixed-effects models on the data from each *Construction* and *Clause* for pairwise comparisons. This analysis showed that the VPE-A condition had significantly longer RTs than the VPE-C condition ($\beta = 0.371$, $SE = 0.098$, $p < .001$). This finding may be attributable to processing difficulty stemming from causal reasoning, as discussed above; it could also have resulted from L1 influence because the closest analogues of VPE in Korean (Argument Ellipsis, *Kulay* 'Do So' Anaphora, and Pseudo-VPE) are generally ungrammatical in adjunct clauses[5] (see §2.1). Another finding was that there was no significant RT difference between the VPE-A condition and the Gapping-A condition ($\beta = -0.164$, $SE = 0.096$, $p = .101$). On the other hand, the Gapping-C condition had significantly longer RTs than the VPE-C condition ($\beta = 0.196$, $SE = 0.093$, $p = .047$) but not the Gapping-A condition ($\beta = -0.058$, $SE = 0.065$, $p = .375$). This indicates that Gapping in conjunct clauses is

[5] Pseudo-VPE is ungrammatical in adjunct clauses. Argument Ellipsis is grammatical in adjunct clauses only when its subject in the ellipsis clause has the nominative case marker (vs. *-to* 'also'), and *Kulay* 'Do So' Anaphora is grammatical in adjunct clauses only when its antecedent precedes it in a separate sentence or main clause.

harder to process and judge than VPE in conjunct clauses but not ungrammatical Gapping in adjunct clauses.

To examine a potential effect of *Proficiency* in the late L2ers' RTs, an additional linear mixed-effects regression model was fitted to their RTs with the factor *Proficiency* added. This model did not find any significant effects related to *Proficiency*, as shown in Table 5.20.

Table 5.20
Results of the Linear Mixed-Effects Regression for the Late L2er Reaction Time Data in Study 2 with the Factor Proficiency *Added*

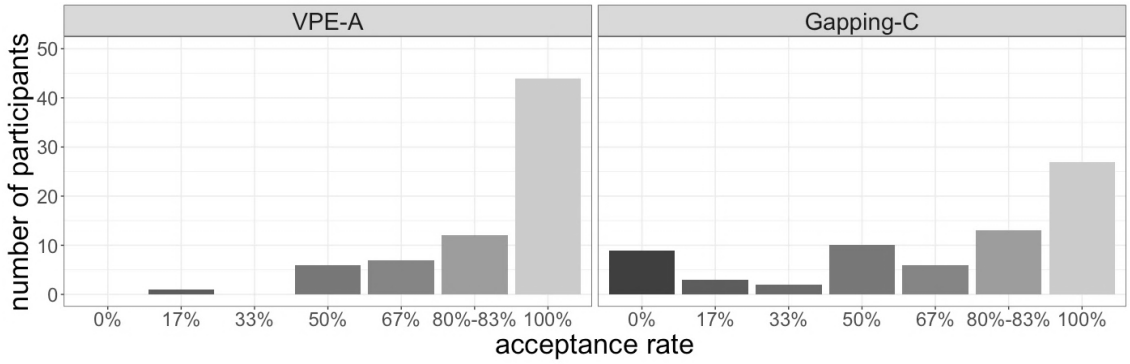
	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.756	0.066	< .001
Construction	0.020	0.080	.806
Clause	0.189	0.070	.012
Proficiency	−0.001	0.026	.955
Construction × Clause	−0.351	0.162	.039
Construction × Proficiency	−0.009	0.034	.785
Clause × Proficiency	−0.041	0.028	.150
Construction × Clause × Proficiency	−0.040	0.058	.498

Note. Model formula: lmer(RT ~ Construction * Clause * Proficiency + (1 + Construction * Clause | participant) + (1 + Construction * Clause | item))

Turning back to the main issue, I have argued that my analysis of the RT data indicates that the lower acceptance rates for the VPE-A and Gapping-C conditions in the L1 adult controls is attributable to processing difficulty. However, as discussed in the beginning of this section, the processing difficulties in these two conditions might be attributable to different sources: Whereas VPE in adjunct clauses requires working out the semantic relationship between the *because*-clause and the clause preceding it, Gapping in conjunct clauses involves a complex structure with a focused remnant (material stranded by the Gapping operation) after the verb gap (e.g., *pasta* in *Sara made pizza, and Kelly [e] pasta.*).

Following this line of reasoning, I next look at whether there was any difference between the two conditions among the L1 adults at the level of the individual. This analysis showed that whereas all but one L1 adult accepted VPE in adjunct clauses at least 50% of the time, Gapping in conjunct clauses displayed a wide range of acceptance rates; this is shown in Figure 5.16.

Figure 5.16
Distribution of L1 Adult Acceptance Rates of VPE-A and Gapping-C



Notably, 20% of the 70 L1 adults controls had less than a 50% acceptance rate for the Gapping-C condition, suggesting that Gapping in conjunct clauses is harder to process than VPE in adjunct clauses and that this processing burden further affected the L1 adults' judgment patterns. This finding gives us reason to think that the relatively lower rates of acceptance for the Gapping-C condition on the part of the L1 children and the L2ers may not necessarily mean that they lack knowledge that Gapping in conjunct clauses is grammatical.

5.4.2 Developmental sequences for the grammaticality contrast between VP-Ellipsis and Gapping.

To review the results: The L1 children and the early L2ers did not uniformly display the target grammaticality contrast between VPE and Gapping, but effects of *Age* and *Proficiency* were observed such that older L1 children and higher-proficiency early L2ers showed evidence of having developed target-like knowledge of VPE and Gapping in English despite the learnability problems they faced.

As a further step toward understanding the developmental sequences involved in acquiring the grammaticality contrast between VPE and Gapping, I inspected the judgments of individual participants from all the learner groups using *Age* as a mediating factor for the L1 children and *Proficiency* as a mediating factor for the L2ers. The four criteria that I used for this investigation are presented in Table 5.21.

Table 5.21
Criteria for Individual Analysis in Study 2

Criterion	Item
A	Knowledge of VPE
B	Knowledge of the ungrammaticality of *backward Gapping
C	Knowledge of the ungrammaticality of *Gapping in adjunct clauses
D	Knowledge of the contrast between Gapping in conjunct clauses and *Gapping in adjunct clauses

That nine L1 adults rejected all six Gapping-C items (presumably due to processing difficulty) leaves us with the question of how to interpret participants' judgments for these sentences. I thus decided to compare the acceptance rates in the Gapping-C and Gapping-A conditions instead of analyzing only the data from the Gapping-C condition (Criterion D). Participants received a check mark (✓) for each criterion they passed, as will be shown in Tables 5.22–5.24.

To ascertain whether a participant possessed target-like knowledge of VPE (Criterion A), I computed the cut-off point at which the accuracy rate becomes significantly higher than chance performance based on a binomial distribution (henceforth, the 'comparison-against-chance diagnostic') following K. Kim (2014). Specifically, I calculated the minimum number of VPE sentences that participants needed to judge correctly in order for their accuracy rate to be significantly above chance performance. Because two of the five response options were correct for each trial (i.e., the two smiling faces), the probability of selecting a correct response at random was always 2/5 (0.4). Based on this value, I further calculated that participants needed to judge at least 8 of the 12 VPE sentences correctly in order for their accuracy to be significantly above chance performance,[6] and so *having 8 or more correct* was set as the criterion for *presence of knowledge*.[7]

Knowledge of the ungrammaticality of backward Gapping was Criterion B. The AJT had 3 fillers targeting the ungrammatical backward Gapping pattern (e.g., *Ryan the chair, and I liked the desk.*), which (at the right level of abstraction) is grammatical in Korean. Because there were

[6] The *p*-value calculated based on a binomial distribution is: {12!/(8!*4!)}*0.4^8*(1−0.4)^4 = 0.042.
[7] Although there was one L1 adult (L1A_53) who accepted only 7 of the VPE sentences, I did not adjust the criterion because the probability of having 7 correct out of 12 items is 0.101, which exceeds the marginal significance level.

only 3 items of this type, I set the criterion at *having all 3 items correct* (probability: 0.064), which reached only marginal significance.[8]

Criterion C was the presence of knowledge that Gapping is ungrammatical in adjunct clauses. Again, the comparison-against-chance diagnostic was used for this criterion, i.e., *having 5 or more correct out of 6 items on the Gapping-A condition* (probability = 0.037).[9, 10]

Criterion D was whether a participant showed the contrast between Gapping-C and Gapping-A. Because even the L1 adults, I argued, had processing difficulty when judging the Gapping-C sentences, I used a less rigorous criterion here according to which the number of correct items for the Gapping-C condition simply needed to be numerically higher than the number of correct items for the Gapping-A condition.[11]

The next step was to group learners together based on the extent of systematicity in their responses so as to allow inferences regarding possible developmental stages. I identified developmental stages based on the following requirements: (a) Each developmental stage should begin with at least two consecutive learners (sorted by *Age* for the L1 children and by *Proficiency* for the L2ers) who pass the criterion at issue and (b) no stage should have two or more consecutive learners who fail to pass the criterion.

There were four distinct developmental stages apparent in the data from the L1 children, as shown in Table 5.22, where they are ordered by ascending *Age*. The first stage, knowledge that VPE is possible, arose by age 5;4. The next stage, knowledge that backward Gapping is ungrammatical, emerged by age 5;10. In the third stage, the L1 children show they know that Gapping is impossible in adjunct clauses by age 6;3. Finally, by age 6;6, they displayed the numerical preference for (possible) Gapping in conjunct clauses over (impossible) Gapping in adjunct clauses.

---

[8] Nine L1 adults failed this criterion, with six accepting backward Gapping once and three more than once.

[9] The *p*-value calculated based on a binomial distribution is: {6!/(5!*1!)}*0.4^5*(1−0.4)^1 = 0.037.

[10] Three L1 adults did not pass this criterion. L1A_19 and L1A_66 accurately judged only 4 out of 6 items. As the probability of having 4 correct out of 6 is 0.138, which does not reach marginal significance, I did not adjust the criterion. L1A_34 got only 2 correct, a result for which I have no explanation.

[11] As mentioned above, there were nine L1 adults who did not accept the Gapping-C sentences at all and so failed to pass this criterion. Still, none of these participants accepted the Gapping-A sentences more often than the Gapping-C sentences.

Table 5.22

*Developmental Sequence of the Grammaticality Contrast between VPE and Gapping: L1 Children in Chronological Order (*n = *33)*

Age	Participant	Knowledge of the grammaticality of VPE	Knowledge of the ungrammaticality of *backward Gapping	Knowledge of the ungrammaticality of *Gapping-A	Knowledge of the contrast between Gapping-C and *Gapping-A	Developmental stage
3;3	L1C_04		✓			
3;11	L1C_05	✓			✓	
4;0	L1C_06			✓	✓	
5;0	L1C_03					
5;4	L1C_10	✓			✓	(1) Knowledge of the grammaticality of VPE
5;6	L1C_28	✓	✓			
5;8	L1C_25	✓				
5;10	L1C_13	✓	✓	✓		(2) Knowledge of the ungrammaticality of *backward Gapping
5;11	L1C_14		✓		✓	
5;11	L1C_26	✓	✓	✓	✓	
6;1	L1C_20	✓				
6;3	L1C_17	✓		✓		(3) Knowledge of the ungrammaticality of *Gapping-A
6;3	L1C_24	✓	✓	✓		
6;6	L1C_31	✓				
6;6	L1C_15	✓	✓	✓	✓	(4) Knowledge of the contrast between Gapping-C and *Gapping-A
6;7	L1C_18	✓		✓	✓	
6;8	L1C_07	✓			✓	
6;8	L1C_12	✓	✓	✓		
6;8	L1C_27	✓			✓	
6;9	L1C_33	✓			✓	
6;9	L1C_23	✓	✓	✓		
6;10	L1C_16	✓			✓	
6;10	L1C_30	✓	✓		✓	
6;11	L1C_09	✓	✓	✓	✓	
6;11	L1C_11		✓	✓	✓	
6;11	L1C_21	✓	✓	✓	✓	
7;1	L1C_22	✓	✓	✓		
7;1	L1C_32	✓	✓	✓		
7;2	L1C_29	✓		✓	✓	
7;2	L1C_19	✓	✓	✓		
7;7	L1C_01	✓	✓	✓	✓	
7;8	L1C_08	✓	✓			
7;9	L1C_02	✓	✓	✓	✓	

Notes. The check mark (✓) indicates presence of knowledge. The four criteria were as follows: knowledge of the grammaticality of VPE (8–12 correct out of 12); knowledge of the ungrammaticality of *backward Gapping (3 correct out of 3); knowledge of the ungrammaticality of *Gapping-A (5–6 correct out of 6); knowledge of the contrast between Gapping-C and *Gapping-A (correct Gapping-C > correct *Gapping-A).

The developmental path for the early L2ers—as shown in Table 5.23, where participants are listed in order of ascending *Proficiency*—differed from that of the L1 children. In the first stage, the early L2ers showed the grammaticality contrast between (possible) Gapping in conjunct clauses and (impossible) Gapping in adjunct clauses. Next, they developed knowledge

of the grammaticality of VPE. The third stage was knowledge of ungrammatical Gapping in adjunct clauses, and the fourth stage was knowledge of ungrammatical backward Gapping. The late emergence of knowledge of the ungrammaticality of backward Gapping may have come from the L2ers' L1 Korean, which allows backward Gapping.

Table 5.23
*Developmental Sequences of the Grammaticality Contrast between VPE and Gapping: Early L2ers in Proficiency Order (*n = *27)*

Proficiency	Participant	Knowledge of the grammaticality of VPE	Knowledge of the ungrammaticality of *backward Gapping	Knowledge of the ungrammaticality of *Gapping-A	Knowledge of the contrast between Gapping-C and *Gapping-A	Developmental stage
−6.65 (L)	EL2_11				✓	(1) Knowledge of the contrast between Gapping-C and *Gapping-A
−5.98 (L)	EL2_04		✓		✓	
−3.80 (L)	EL2_23	✓			✓	
−3.41 (L)	EL2_12		✓	✓		
−3.25 (L)	EL2_24				✓	
−2.65 (L)	EL2_05				✓	
−2.52 (L)	EL2_21	✓	✓	✓		
−2.49 (L)	EL2_10					
−2.36 (L)	EL2_07			✓	✓	
−2.35 (L)	EL2_08	✓				(2) Knowledge of the grammaticality of VPE
−1.96 (L)	EL2_27	✓			✓	
−1.10 (M)	EL2_18				✓	
−0.78 (M)	EL2_25			✓	✓	(3) Knowledge of the ungrammaticality of *Gapping-A
−0.71 (M)	EL2_15	✓	✓	✓	✓	(4) Knowledge of the ungrammaticality of *backward Gapping
−0.19 (M)	EL2_02	✓	✓	✓		
0.39 (M)	EL2_03	✓		✓	✓	
0.51 (M)	EL2_06	✓	✓	✓	✓	
0.51 (M)	EL2_26		✓		✓	
0.53 (M)	EL2_01	✓		✓	✓	
0.80 (M)	EL2_22	✓	✓	✓	✓	
1.45 (H)	EL2_09			✓	✓	
1.62 (H)	EL2_13	✓		✓		
1.69 (H)	EL2_19	✓	✓	✓		
1.90 (H)	EL2_17	✓	✓	✓	✓	
2.79 (H)	EL2_20	✓	✓	✓		
2.90 (H)	EL2_16	✓	✓	✓		
3.51 (H)	EL2_14	✓			✓	

Notes. (L), (M), and (H) respectively indicate the Lower proficiency group, the Medium proficiency group, and the Higher proficiency group. The check mark (✓) indicates presence of knowledge. The four criteria were as follows: knowledge of the grammaticality of VPE (8–12 correct out of 12); knowledge of the ungrammaticality of *backward Gapping (3 correct out of 3); knowledge of the ungrammaticality of *Gapping-A (5–6 correct out of 6); knowledge of the contrast between Gapping-C and *Gapping-A (correct Gapping-C > correct *Gapping-A).

As for the late L2ers, most of them passed all four criteria, as shown in Table 5.24. This result can be attributed to their (relatively) high level of English proficiency; the late L2ers' proficiency scores were significantly higher than the early L2ers ($t(55) = 2.548$, $p = .014$, Cohen's $d = 0.690$; see also Appendix A).

Table 5.24
*Developmental Sequence of the Grammaticality Contrast between VPE and Gapping: Late L2ers in Proficiency Order (*n = *30)*

Proficiency	Participant	Knowledge of the grammaticality of VPE	Knowledge of the ungrammaticality of *backward Gapping	Knowledge of the ungrammaticality of *Gapping-A	Knowledge of the contrast between Gapping-C and *Gapping-A	Developmental stage
−4.36 (L)	LL2_26			✓	✓	NA
−2.31 (L)	LL2_22	✓	✓	✓	✓	
−2.00 (L)	LL2_05	✓		✓	✓	
−1.36 (L)	LL2_24	✓		✓		
−0.79 (M)	LL2_18	✓	✓	✓	✓	
−0.31 (M)	LL2_08	✓	✓	✓		
−0.24 (M)	LL2_19	✓	✓	✓	✓	
0.14 (M)	LL2_01	✓	✓	✓	✓	
0.14 (M)	LL2_20	✓	✓	✓	✓	
0.22 (M)	LL2_16	✓	✓	✓	✓	
0.24 (M)	LL2_14	✓	✓	✓	✓	
0.25 (M)	LL2_04	✓	✓	✓		
0.46 (M)	LL2_11	✓	✓	✓	✓	
0.59 (M)	LL2_21	✓	✓	✓		
0.64 (M)	LL2_10	✓		✓		
0.80 (M)	LL2_27	✓	✓	✓	✓	
0.72 (M)	LL2_28	✓	✓	✓	✓	
0.90 (M)	LL2_02	✓		✓		
1.44 (H)	LL2_23	✓	✓	✓		
1.44 (H)	LL2_06	✓	✓	✓	✓	
1.52 (H)	LL2_13	✓	✓	✓	✓	
1.85 (H)	LL2_17	✓	✓	✓	✓	
2.09 (H)	LL2_07	✓	✓	✓	✓	
2.16 (H)	LL2_15	✓	✓	✓	✓	
2.39 (H)	LL2_09	✓	✓	✓	✓	
2.39 (H)	LL2_29	✓	✓	✓	✓	
2.41 (H)	LL2_25	✓	✓	✓	✓	
2.80 (H)	LL2_03	✓		✓		
3.53 (H)	LL2_12	✓	✓	✓	✓	
3.84 (H)	LL2_30	✓	✓	✓	✓	

Notes. (L), (M), and (H) respectively indicate the Lower proficiency group, the Medium proficiency group, and the Higher proficiency group. The check mark (✓) indicates presence of knowledge. The four criteria were as follows: knowledge of the grammaticality of VPE (8–12 correct out of 12); knowledge of the ungrammaticality of *backward Gapping (3 correct out of 3); knowledge of the ungrammaticality of *Gapping-A (5–6 correct out of 6); knowledge of the contrast between Gapping-C and *Gapping-A (correct Gapping-C > correct *Gapping-A).

It should be noted that the ungrammatical backward Gapping pattern was consistently rejected by the older L1 children,[12] the higher-proficiency early L2ers,[13] and most of the late L2ers.[14] However, the same participants accepted the grammatical forward Gapping pattern only 28–44% of the time (L1 children: $M = 33.70\%$, $SD = 34.42\%$; early L2ers: $M = 28.52\%$, $SD = 31.41\%$; late L2ers: $M = 43.94\%$, $SD = 42.27\%$), just like the L1-Japanese L2ers of English in O'Grady's (1999) study. From this, it might be tempting to conclude that these participants lack knowledge that Gapping is grammatical in English.

However, all three learner groups accepted grammatical Gapping in conjunct clauses more often than ungrammatical Gapping in adjunct clauses (all ps $< .05$; see §5.3). Furthermore, the L1 children ($\beta = -2.263$, $SE = 0.578$, $p = .029$) and the late L2ers ($\beta = -4.532$, $SE = 1.638$, $p = .006$) both accepted grammatical Gapping in conjunct clauses more often than ungrammatical backward Gapping (cf. L1 adults: $\beta = -10.275$, $SE = 1.646$, $p < .001$); although the early L2ers as a group did not show a significant difference between grammatical Gapping in conjunct clauses and ungrammatical backward Gapping ($\beta = -0.734$, $SE = 0.488$, $p = .133$), the difference approached statistical significance when only those in the Medium and Higher proficiency groups were included in the analysis ($\beta = -0.879$, $SE = 0.465$, $p = .059$). In addition, passing the comparison-against-chance diagnostic for the Gapping-C condition (and for the Gapping-A condition) was not impossible for all the L1 children and L2ers: The oldest L1 child (L1C_02, aged 7;9) and six of the late L2ers (LL2_06, LL2_07, LL2_17, LL2_18, LL2_20, LL2_22) got five or more correct out of six Gapping-C items. I thus argue that the low acceptance rates of the Gapping-C condition stem from processing difficulty and not from lack of grammatical knowledge.

5.4.3 Limitations and future directions.

One interesting and curious finding, for which I lack an explanation, is that the older L1 children, the higher-proficiency early L2ers, and the higher-proficiency late L2ers were all more likely than the other participants in their groups to reject the grammatical Gapping-C condition

[12] L1C_09, L1C_11, L1C_21, L1C_22, L1C_32, L1C_19, L1C_01, L1C_08, and L1C_02
[13] EL2_15, EL2_02, EL2_06, EL2_26, EL2_22, EL2_19, EL2_17, EL2_20, and EL2_16
[14] LL2_26, LL2_22, LL2_05, LL2_18, LL2_19, LL2_01, LL2_20, LL2_16, LL2_14, LL2_11, LL2_27, LL2_28, LL2_06, LL2_13, LL2_17, LL2_07, LL2_15, LL2_09, LL2_29, LL2_25, LL2_12, and LL2_30

(see Figures 5.7, 5.13, and 5.14). It remains to be seen whether reducing processing difficulty would boost acceptance rates in the Gapping-C condition.

Future researchers could attempt to lighten Gapping-C processing load by adding a preceding context that is felicitous for Gapping or by using stimuli with PPs following the verb gaps (e.g., *Mom sat on the sofa, and Dad [e] on the chair.*). If implementing such measures were to lead to greater acceptance of Gapping in conjunct clauses, it would provide further support for the position that processing pressures were responsible for the relatively lower acceptance rates for Gapping in conjunct clauses observed in this study.

5.5 Conclusion

In sum, this study showed that older L1 children and higher-proficiency L2ers were able to acquire target knowledge of the grammaticality contrast between VPE and Gapping in English. The L1 children displayed the contrast as early as age 5;11. Six out of the total 33 children showed clear evidence of having acquired the contrast. For the early L2ers, the results showed that the more proficient participants were more likely to exhibit target-like performance; there were 4 early L2ers (out of 27) who evinced a contrast between VPE and Gapping. Twenty of the 30 late L2ers showed the target contrast, too. These results indicate that L1 children and L1-Korean L2ers are indeed able to overcome the learnability problems involved in acquiring the grammaticality contrast between VPE and Gapping.

CHAPTER VI
STUDY 3: L1 AND L2 ACQUISITION OF THE INTERPRETATION CONTRAST BETWEEN VP-ELLIPSIS AND GAPPING IN ENGLISH

This chapter describes a study investigating whether L1-English children and L1-Korean early and late L2ers of English can develop the interpretation contrasts between VP-Ellipsis (VPE) and Gapping. In §6.1, I provide the research questions. Then I lay out the method of the main task, i.e., the picture-sentence matching task, in §6.2 and report the judgment data from this task in §6.3. This is followed by the discussion of these data in §6.4, where reaction time data are also examined. Section 6.5 concludes this chapter.

6.1 Research Questions

In this study, I investigated the following research questions:

(a) How early do L1-English-acquiring children know the contrast between possible vs. impossible interpretations of VPE and Gapping in English?
(b) Do early and late L1-Korean L2ers of English come to know the contrast between possible vs. impossible interpretations of VPE and Gapping in English? What role does L2 proficiency play?

6.2 Method

6.2.1 Participants.

There were four groups of participants in this study: 32 L1 adults (from Study 2, L1A_39–L1A_70), 24 L1 children (from Study 2, L1C_10–L1C_33), 27 early L2ers (from Study 2, EL2_01–EL2_27), and 30 late L2ers (from Study 2, LL2_01–LL2_30). The proficiency scores for these participants are provided in Appendix A. Note that they all took part in Study 2 (see Chapter 5), which had 70 L1 adults (L1A_01–L1A_70), 33 L1 children (L1C_01–L1C_33), and the same 27 early L2ers and 30 late L2ers. The L1 children group included 5-year-olds (n = 6), 6-year-olds (n = 14), and 7-year-olds (n = 4). The background information for the four groups is summarized in Table 6.1.

Table 6.1
Background Information for Participants in Study 3

	Age at testing	Age of English onset	Length of residence in an English-speaking country in months
L1 Adults (*n* = 32)	23.78 (*SD* = 5.76; range = 19–49)	N/A	N/A
L1 Children (*n* = 24)	5.92 (*SD* = 0.65; range = 5–7)	N/A	N/A
Early L2ers (*n* = 27)	8.52 (*SD* = 1.63; range = 5–12)	4.96 (*SD* = 0.76; range = 4–6)	1.15 (*SD* = 3.46; range = 0–13)
Late L2ers (*n* = 30)	23.03 (*SD* = 2.92; range = 18–30)	8.83 (*SD* = 1.09; range = 8–12)	2.83 (*SD* = 5.98; range = 0–24)

6.2.2 Picture-sentence matching task.

6.2.2.1 Materials.

The picture-sentence matching task crossed the factors *Construction* (VPE; Gapping) and *Interpretation* (subject reading (SR); object reading (OR)) in a 2 × 2 Latin square design, as laid out in Table 6.2. The critical condition was VPE-OR (impossible in English, possible in Korean[1]). The sentence stimuli were distributed across four lists, each with 16 critical items (k = 4 per condition) alongside 8 fillers. They were recorded by an English native speaker (with phonology training) in natural prosody, i.e., the prosody exemplified in Carlson (2001) for the Gapping-SR and Gapping-OR sentences. Appendix G provides the prosody of a sample stimulus for each critical condition using a pitch track. The length of the stimuli (including the fillers) ranged from 6 to 10 words. To minimize the processing burden on participants, frequent and easy words were used for the stimuli.

[1] Korean does not have VPE. But even the three false Korean analogues of VPE (Argument Ellipsis, *Kulay* 'Do So' Anaphora, Pseudo-VPE) all allow both SRs and ORs (see §2.1).

Table 6.2
Sample Stimuli of the Critical Conditions in the Picture-Sentence Matching Task

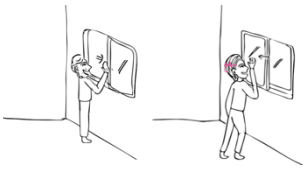
Condition	Pictures	Audio stimuli
VPE-SR (MATCH; $k = 4$)		Story: Mom hugged the boy at home. Dad hugged the boy at home too. Target sentence: Mom hugged the boy at home and Dad did too.
*VPE-OR (MISMATCH; $k = 4$)		Story: Mom hugged the boy at home. Mom hugged Dad at home too. Target sentence: *Mom hugged the boy at home and Dad did too.
Gapping-SR (MATCH; $k = 4$)		Story: Mom hugged the boy at home. Dad hugged the boy in the park. Target sentence: Mom hugged the boy at home and Dad in the park.
Gapping-OR (MATCH; $k = 4$)		Story: Mom hugged the boy at home. Mom hugged Dad in the park. Target sentence: Mom hugged the boy at home and Dad in the park.

There were two types of fillers: One type involved a verb mismatch, as shown in Table 6.3, and the other involved an object mismatch, as shown in Table 6.4. Because the target answer for 12 out of 16 critical items was "MATCH," the target answer for all 8 fillers was "MISMATCH" to ensure that there would be equal numbers of "MATCH" and "MISMATCH" stimuli across the task as a whole. See Appendix H for the full list of experimental items (including the fillers) and Appendix I for the results of the filler items for each group.

Table 6.3
*Sample Stimulus of the *Verb Mismatch Fillers in the Picture-Sentence Matching Task*

Condition	Pictures	Audio stimuli
*Verb mismatch (MISMATCH; $k = 4$)		Story: The woman liked yellow flowers. The man hated red flowers. Target sentence: *The woman liked yellow flowers and the man red flowers.

Table 6.4
*Sample Stimulus of the *Object Mismatch Fillers in the Picture-Sentence Matching Task*

Condition	Pictures	Audio stimuli
*Object mismatch (MISMATCH; $k = 4$)		Story: Kyle opened the window. Helen closed the window. Target sentence: *Kyle opened the window and Helen closed the door.

6.2.2.2 Procedure.

Participants completed the picture-sentence matching task in PsychoPy (Peirce, 2017) after finishing the language background questionnaire (see Appendix E) and the acceptability judgment task but before starting the picture narration proficiency task (see §5.2.2.2). The task began with the instructions and two practice items that were unrelated to the other stimuli. The participants then completed the target and filler items, which were presented in pseudo-random order. For each item, participants were first given two pictures accompanied by a simple audio narration in a child-friendly voice (see Tables 6.2–6.4). Next, a (female) puppet popped up on the screen and the pre-recorded (male) narrator asked her to say what was happening in the pictures. The puppet then produced the target sentence. Finally, the participants judged whether the target sentence matched the pictures or not by pressing one of three buttons on the keyboard: a smiling face, a frowning face, or a question mark indicating 'I don't know' (see Figure 6.1). In addition to their judgments, which were my focus of analysis, their reaction times (RTs) were recorded from the offset of the audio stimulus; they were expected to provide more nuanced information about participants' judgments. Completing this task took approximately 20 minutes.

Figure 6.1
Judgment Options for the Picture-Sentence Matching Task

Note. Adapted from Ambridge et al. (2008).

6.2.3 Data analysis.

There were two types of data obtained from the picture-sentence matching task: judgment data and RT data. All 'I don't know' responses were removed prior to analysis, which affected

0.39% of the L1 adult data, 2.43% of the L1 child data, 1.23% of the early L2 data, 0.83% of the late L2 data.

For analysis of the judgment data, smiling face responses were recoded as "1" (accept), and frowning face responses were recoded as "0" (reject). The responses for each group were then analyzed by means of a logistic mixed-effects regression model with *Construction* and *Interpretation* as binary fixed effects (contrast-coded with [−.5, .5]) and *participant* and *item* as random effects. To investigate a potential effect of *Age* in the L1 children and *Proficiency* in the L2ers, a separate logistic mixed-effects regression model was constructed for each learner group, with *Construction* and *Interpretation* as binary fixed effects (contrast-coded with [−.5, .5]), *Age* (L1 child data only) and *Proficiency* (L2 data only) as continuous fixed effects, and *participant* and *item* as random effects.

Data trimming for the RTs was conducted for each group. I began with the removal of extreme values that were longer than 30 seconds in duration, which affected 0.39% of the L1 adult data, 0.27% of the L1 child data, 0.00% of the early L2 data, and 3.37% of the late L2 data. RTs that were more than 1.5 standard deviations above or below the condition either by participant or by item were removed; the by-participant removal procedure affected 6.88% of the L1 adult data, 5.32% of the L1 child data, 4.46% of the early L2 data, 7.41% of the late L2 data; the by-item removal procedure affected 8.23% of the L1 adult data, 9.83% of the L1 child data, 9.34% of the early L2 data, and 8.71% of the late L2 data. The RT data from one late L2er (LL2_27) were lost due to technical difficulties. A linear mixed-effects model was fitted to the trimmed RTs for each group with *Construction* and *Clause* as binary fixed effects (contrast-coded with [−.5, .5]) and *participant* and *item* as random effects. The L1 child data, the early L2 data, and the late L2 data were further investigated for potential effects of *Age* (for the L1 children) and *Proficiency* (for the L2ers); separate linear mixed-effects models were constructed with either *Age* (L1 child data only) or *Proficiency* (L2 data only) as a continuous fixed effect, *Construction* and *Clause* as binary fixed effects (contrast-coded with [−.5, .5]), and *participant* and *item* as random effects.

All mixed-effects models were constructed with the maximal random effects structure permitted by the design (Barr et al., 2013; Matuschek et al., 2017; Stroup, 2012) in R (R Core Team, 2018). The model formula for each analysis is reported in each corresponding results table.

6.3 Results

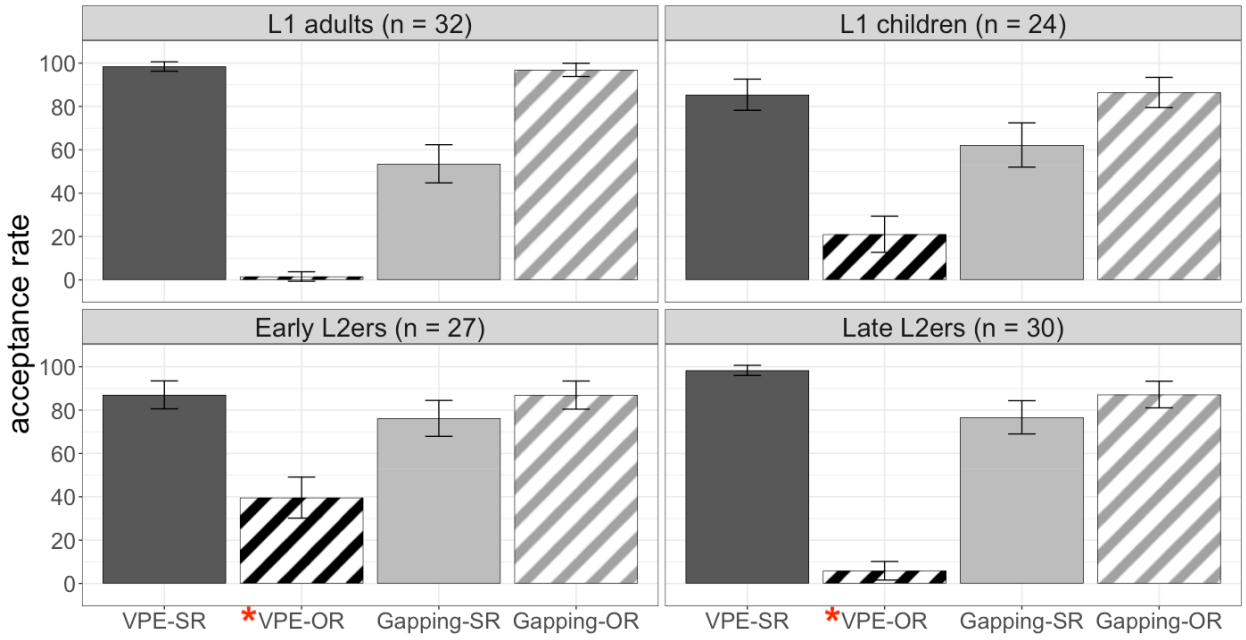
The means (and standard deviations) of each group's acceptance rates for the four conditions are presented in Table 6.5 and Figure 6.2.

Table 6.5
Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Group

	VPE-SR	***VPE-OR**	**Gapping-SR**	**Gapping-OR**
L1 Adults	98.44	1.56	53.54	96.88
(*n* = 32)	(*SD* = 12.45)	(*SD* = 12.45)	(*SD* = 50.07)	(*SD* = 17.47)
L1 Children	85.42	21.05	62.22	86.46
(*n* = 24)	(*SD* = 35.48)	(*SD* = 40.98)	(*SD* = 48.75)	(*SD* = 34.40)
Early L2ers	87.04	39.62	76.19	86.92
(*n* = 27)	(*SD* = 33.75)	(*SD* = 49.14)	(*SD* = 42.80)	(*SD* = 33.88)
Late L2ers	98.32	5.88	76.67	87.18
(*n* = 30)	(*SD* = 12.91)	(*SD* = 23.63)	(*SD* = 42.47)	(*SD* = 33.58)

Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading.

Figure 6.2
Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Group



Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading. Error bars represent 95% confidence intervals.

As predicted, L1 adult controls accepted the VPE-SR and Gapping-OR conditions but rejected the VPE-OR condition. However, their acceptance rate for the Gapping-SR condition was only slightly above 50% with a large standard deviation—a result that is difficult to interpret. I argue that this pattern is attributable to processing difficulty (see §6.4.1). While the acceptance patterns of the L1 children, early L2ers, and late L2ers resembled those of the L1 adults, there were slight differences from one group to the next. For example, all the learner groups exhibited higher acceptance rates for the Gapping-SR condition than the L1 adult group did. Also, the L1 children and the early L2ers accepted the VPE-OR condition (i.e., the critical condition in this study) at much higher rates than the L1 adults did.

In the following sections, I report the results of additional analyses carried out on the data from the picture-sentence matching task. One reason to perform these analyses was to explore each group's response patterns in more detail. Another reason was to investigate the possibility that the relatively high acceptance rates in the (impossible) VPE-OR condition in the data from the L1 children and the early L2ers is an artifact of group averaging and that some of the participants in these groups did in fact possess target knowledge that VPE with an object reading is impossible in English.

6.3.1 L1 adults.

The logistic mixed-effects model for L1 adults showed a marginally significant effect of *Interpretation* ($\beta = 5.605$, $SE = 2.911$, $p = .054$), with higher acceptance rates for the SR conditions than the OR conditions, but no effect of *Construction* ($\beta = -1.576$, $SE = 2.746$, $p = .566$), as shown in Table 6.6.

Table 6.6
Results of the Logistic Mixed-Effects Regression for the L1 Adult Judgment Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.731	1.435	.228
Construction	−1.576	2.746	.566
Interpretation	5.605	2.911	.054
Construction × Interpretation	20.108	5.756	< .001

Note. Model formula: glmer(Judgment ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 | item))

This group also showed a significant interaction between *Construction* and *Interpretation* (β = 20.108, *SE* = 5.756, $p < .001$). To unpack this interaction, I constructed separate mixed-effects models for pairwise comparisons. These models found that L1 adults accepted the Gapping-SR condition significantly less often than both the VPE-SR condition (β = 9.433, *SE* = 3.284, $p = .004$) and the Gapping-OR condition ($\beta = -3.216$, *SE* = 1.105, $p = .004$). This result may be associated with processing difficulty in the Gapping-SR condition, as will be discussed in §6.4.1. More importantly, L1 adults accepted the VPE-OR condition significantly less often than both the VPE-SR condition (β = 66.102, *SE* = 13.035, $p = .001$) and the Gapping-OR condition ($\beta = -33.074$, *SE* = 7.069, $p < .001$), indicating that they know VPE with an object reading is ungrammatical. In sum, L1 adult controls have the target interpretive ambiguity contrasts between VPE and Gapping.

6.3.2 L1 children.

The logistic mixed-effects model for L1 children revealed a significant effect of *Construction* ($\beta = -1.850$, *SE* = 0.696, $p = .008$) and *Interpretation* (β = 1.621, *SE* = 0.670, $p = .015$), as shown in Table 6.7.

Table 6.7
Results of the Logistic Mixed-Effects Regression for the L1 Child Judgment Data in Study 3

	Estimate (β)	Standard Error (SE)	*p*
(Intercept)	1.002	0.427	.019
Construction	−1.850	0.696	.008
Interpretation	1.621	0.670	.015
Construction × Interpretation	6.250	1.444	< .001

Note. Model formula: glmer(Judgment ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Crucially for the current study, there was a significant interaction between *Construction* and *Interpretation* (β = 6.250, *SE* = 1.444, $p < .001$). To identify the source of the interaction effect, I ran separate mixed-effects models for pairwise comparisons. These models found that L1 children's acceptance rate was significantly lower for the VPE-OR condition than for both the VPE-SR condition (β = 4.556, *SE* = 1.015, $p < .001$) and the Gapping-OR condition ($\beta = -4.738$, *SE* = 0.960, $p < .001$). This indicates that L1 children know that VPE with an object reading is ungrammatical. There was also a marginally significant difference between the Gapping-SR and

Gapping-OR conditions ($\beta = -1.386$, $SE = 0.837$, $p = .098$) such that the former was accepted less often than the latter. This result is attributable to the processing difficulty involved in the Gapping-SR condition, as will be discussed in §6.4.1. However, there was no significant difference between the VPE-SR and Gapping-SR conditions ($\beta = 1.120$, $SE = 0.915$, $p = .221$).

To examine a potential effect of *Age* in the data from the L1 children, a separate logistic mixed-effects regression model was constructed with the factor *Age* added. This model did not reveal a significant effect of *Age* ($\beta = -0.458$, $SE = 0.637$, $p = .473$), as shown in Table 6.8.

Table 6.8
Results of the Logistic Mixed-Effects Regression for the L1 Child Judgment Data in Study 3 with the Factor Age *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	3.982	4.139	.336
Construction	−1.646	5.838	.778
Interpretation	7.415	5.555	.182
Age	−0.458	0.637	.473
Construction × Interpretation	−25.301	11.819	.032
Construction × Age	−0.019	0.896	.983
Interpretation × Age	−0.894	0.853	.295
Construction × Interpretation × Age	4.837	1.818	.008

Note. Model formula: glmer(Judgment ~ Construction * Interpretation * Age + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

In addition, there was no significant interaction between *Construction* and *Age* ($\beta = -0.019$, $SE = 0.896$, $p = .983$) or between *Interpretation* and *Age* ($\beta = -0.894$, $SE = 0.853$, $p = .295$). However, there was a three-way interaction among *Construction*, *Interpretation*, and *Age* ($\beta = 4.837$, $SE = 1.818$, $p = .008$). To investigate the way *Age* affects the interaction between *Construction* and *Interpretation*, I conducted a simple regression analysis with *Age* as the independent variable and the strength of sensitivity to the target interpretation contrasts as the dependent variable. I generated a sensitivity score for each participant by subtracting the mean acceptance rate for the ungrammatical interpretation (VPE-OR) from the average of the mean acceptance rates for the grammatical interpretations (VPE-SR, Gapping-SR, Gapping-OR) as laid out in (1). All participants' sensitivity scores are provided in Appendix J.

(1) Interpretation contrast sensitivity scores

([Mean acceptance rate for VPE-SR] + [Mean acceptance rate for Gapping-SR] + [Mean acceptance rate for Gapping-OR]) / 3 − [Mean acceptance rate for *VPE-OR]

The sensitivity scores obtained from the formula in (1) can be interpreted in essentially the same way as the grammaticality contrast sensitivity scores in Study 2 (see §5.3.2), with higher scores indicating stronger sensitivity to the target contrast. The scores can range from −100 to 100.

The simple regression analysis performed on the sensitivity scores, however, did not show a significant effect of *Age* (β = 15.908, *SE* = 12.113, p = .203). In addition, there was no significant effect of *Age* in the VPE-SR condition (β = 4.961, *SE* = 7.538, p = .517), the VPE-OR condition (β = −17.036, *SE* = 10.862, p = .131), the Gapping-SR condition (β = −21.799, *SE* = 13.957, p = .133), or the Gapping-OR condition (β = 13.455, *SE* = 8.681, p = .135).

To further explore *Age* effects in the L1 child data, I conducted a descriptive statistical analysis with three age groups. Table 6.9 and Figure 6.3 show the results of this analysis.

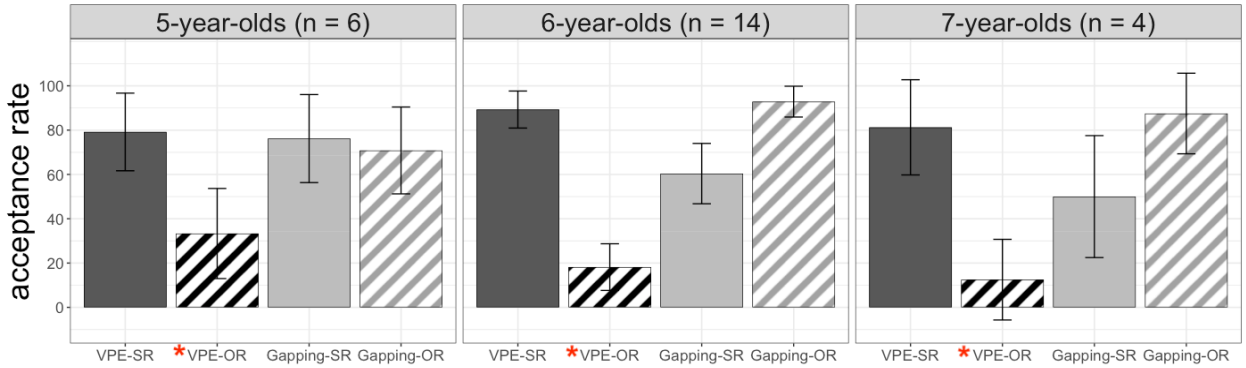
Table 6.9
L1 Children's Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Age Group

	VPE-SR	***VPE-OR**	**Gapping-SR**	**Gapping-OR**
5-year-olds (n = 6)	79.17 (*SD* = 41.49)	33.33 (*SD* = 48.15)	76.19 (*SD* = 43.64)	70.83 (*SD* = 46.43)
6-year-olds (n = 14)	89.29 (*SD* = 31.21)	18.18 (*SD* = 38.92)	60.38 (*SD* = 49.38)	92.86 (*SD* = 25.99)
7-year-olds (n = 4)	81.25 (*SD* = 40.31)	12.50 (*SD* = 34.16)	50.00 (*SD* = 51.64)	87.50 (*SD* = 34.16)

Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading.

Figure 6.3

L1 Children's Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Age Group



Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading. Error bars represent 95% confidence intervals.

The 5-year-olds displayed the target contrast between VPE and Gapping, with uniformly high rates of acceptance for all conditions except the impossible VPE-OR condition. However, their acceptance rates in the VPE-OR condition were still relatively high at 33.33%. The 6-year-olds and 7-year-olds were more target-like, accepting the VPE-OR condition at much lower rates. It should be noted, though, that their acceptance rate for the Gapping-SR condition was not as high as those for the other grammatical conditions. I will return to this point in §6.4.1.

Overall, L1 children had the target interpretation contrast between VPE and Gapping by age 6 at the latest.

6.3.3 Early L2ers.

As shown in Table 6.10, the logistic mixed-effects model for the early L2ers revealed a main effect of *Interpretation* ($\beta = 3.311$, $SE = 1.618$, $p = .041$), with higher acceptance rates for the SR conditions than for the OR conditions, but no effect of *Construction* ($\beta = 0.954$, $SE = 1.498$, $p = .524$).

Table 6.10
Results of the Logistic Mixed-Effects Regression for the Early L2er Judgment Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.583	0.872	.003
Construction	0.954	1.498	.524
Interpretation	3.311	1.618	.041
Construction × Interpretation	8.786	3.183	.006

Note. Model formula: glmer(Judgment ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Importantly, a significant interaction emerged between *Construction* and *Interpretation* (β = 8.786, *SE* = 3.183, p = .006). This interaction was further investigated with planned pairwise comparison analyses. These analyses revealed a significantly lower acceptance rate for the VPE-OR condition than for both the VPE-SR condition (β = 3.353, *SE* = 0.521, p < .001) and the Gapping-OR condition (β = −3.646, *SE* = 0.901, p < .001). This result indicates that the early L2ers knew that VPE cannot have an object reading. In addition, the L1 children's acceptance rate for the Gapping-SR condition was lower than the VPE-SR condition (β = 4.474, *SE* = 2.585, p = .083) and the Gapping-OR condition (β = −1.317, *SE* = 0.799, p = .099), albeit marginally significant. This may be related to processing difficulty involved in the Gapping-SR condition, as will be discussed in §6.4.1.

An additional model constructed to investigate the effect of *Proficiency* did not find a significant effect of *Proficiency* (β = 0.155, *SE* = 0.194, p = .423), as shown in Table 6.11.

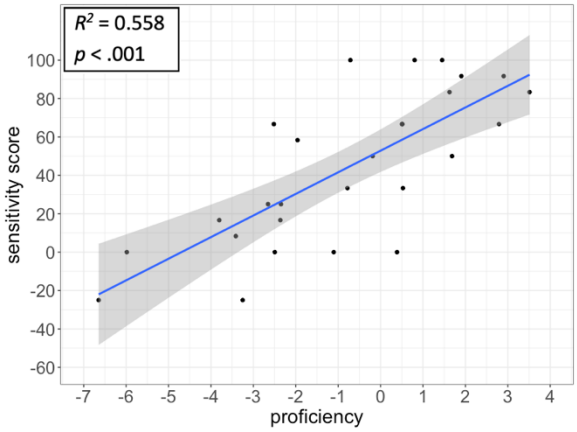
Table 6.11
Results of the Logistic Mixed-Effects Regression for the Early L2er Judgment Data in Study 3 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.540	0.797	.001
Construction	0.311	1.373	.821
Interpretation	3.158	1.451	.029
Proficiency	0.155	0.194	.423
Construction × Interpretation	9.478	2.822	.001
Construction × Proficiency	−0.160	0.271	.555
Interpretation × Proficiency	0.522	0.318	.100
Construction × Interpretation × Proficiency	1.644	0.559	.003

Note. Model formula: glmer(Judgment ~ Construction * Interpretation * Proficiency + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Nor was there a significant interaction between *Construction* and *Proficiency* ($\beta = -0.160$, $SE = 0.271$, $p = .555$) or between *Interpretation* and *Proficiency* ($\beta = 0.522$, $SE = 0.318$, $p = .100$). Crucially, the model for the early L2ers showed a significant three-way interaction among *Construction*, *Interpretation*, and *Proficiency* ($\beta = 1.644$, $SE = 0.559$, $p = .003$). To inspect the modulating role of *Proficiency* in the early L2ers' interpretation contrasts, I ran a simple regression analysis with *Proficiency* as the independent variable and sensitivity scores as the dependent variable (for individual sensitivity scores in the early L2 group, see Appendix J). This analysis revealed a significant effect of *Proficiency* ($\beta = 11.257$, $SE = 2.005$, $p < .001$), as shown in Figure 6.4. This result indicates that as their proficiency increased, the early L2ers had higher sensitivity to the target interpretation contrast between VPE and Gapping.

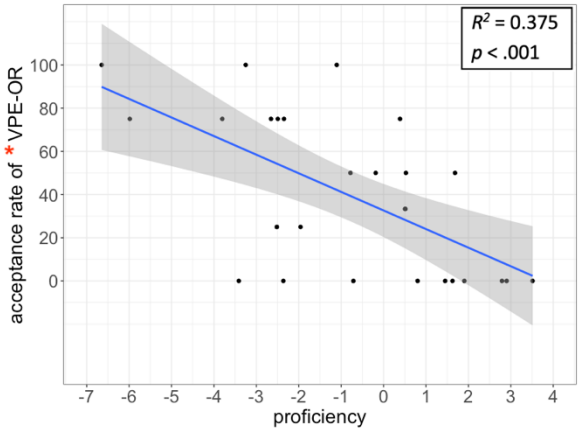
Figure 6.4
Relation between Proficiency and Sensitivity Score for the Early L2ers in Study 3



Note. The shaded region shows the 95% confidence interval for the sample mean.

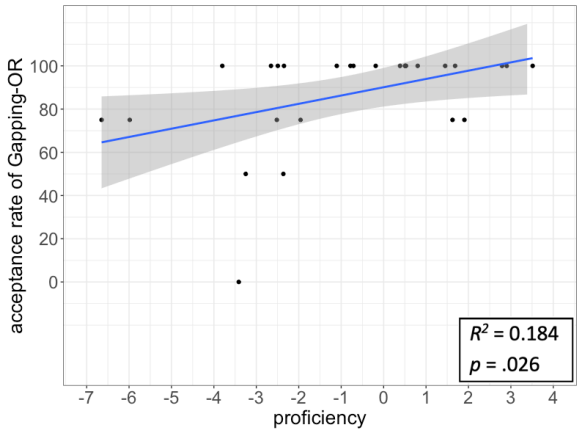
Furthermore, this effect of *Proficiency* in the early L2ers' interpretation contrasts was found to be related to the significant effect of *Proficiency* in the VPE-OR condition ($\beta = -8.610$, $SE = 2.221$, $p < .001$) and in the Gapping-OR condition ($\beta = 3.836$, $SE = 1.616$, $p = .026$). As their *Proficiency* increased, learners were more able to reject VPE-OR and accept Gapping-OR, as shown in Figures 6.5 and 6.6.

Figure 6.5
*Relation between Proficiency and Acceptance Rate of *VPE-OR for the Early L2ers in Study 3*



Note. The shaded region shows the 95% confidence interval for the sample mean.

Figure 6.6
Relation between Proficiency and Acceptance Rate of Gapping-OR for the Early L2ers in Study 3



Note. The shaded region shows the 95% confidence interval for the sample mean.

However, no *Proficiency* effect emerged for the VPE-SR condition (β = 3.220, *SE* = 2.115, p = .140) or the Gapping-SR condition (β = 0.885, *SE* = 2.274, p = .700).

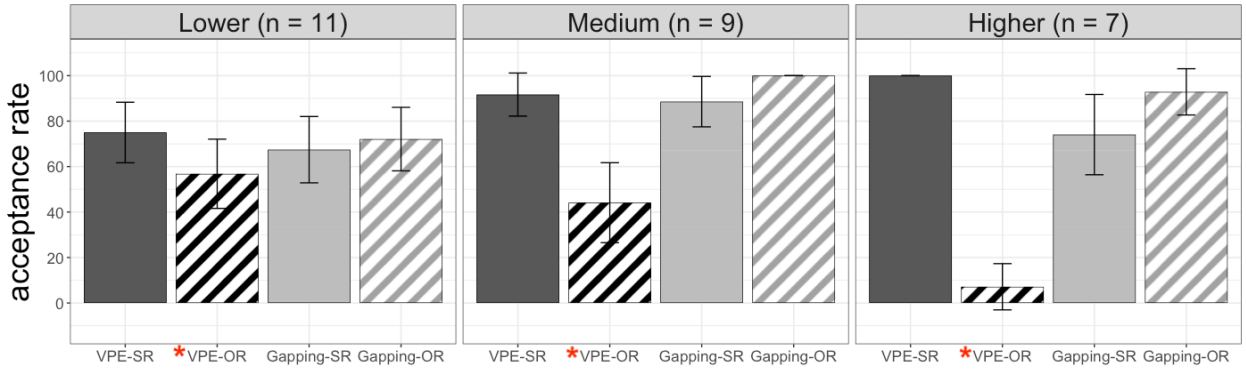
In order to further probe the proficiency effect in the early L2ers, I analyzed their judgment data after dividing them into three proficiency groups based on their proficiency scores (see §5.2.3). The results of this analysis are presented in Table 6.12 and Figure 6.7.

Table 6.12
Early L2ers' Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Proficiency Group

	VPE-SR	***VPE-OR**	**Gapping-SR**	**Gapping-OR**
Lower	75.00	56.82	67.44	72.09
(*n* = 11)	(*SD* = 43.80)	(*SD* = 50.11)	(*SD* = 47.41)	(*SD* = 45.39)
Medium	91.67	44.12	88.57	100.00
(*n* = 9)	(*SD* = 28.03)	(*SD* = 50.40)	(*SD* = 32.28)	(*SD* = 0.00)
Higher	100.00	7.14	74.07	92.86
(*n* = 7)	(*SD* = 0.00)	(*SD* = 26.23)	(*SD* = 44.66)	(*SD* = 26.23)

Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading.

Figure 6.7
Early L2ers' Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Proficiency Group



Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading. Error bars represent 95% confidence intervals.

Although the Lower group and the Medium group had a lower acceptance rate in the ungrammatical VPE-OR condition than in the grammatical conditions, their acceptance rate for the VPE-OR condition was still relatively high at 50%. The Higher group performed very much like the L1 adult controls, successfully rejecting the VPE-OR condition while accepting the other three conditions.

To summarize, proficiency effects were observed in the data from the early L2ers such that learners with higher proficiency were more likely to display target-like knowledge of the interpretation contrast between VPE and Gapping.

6.3.4 Late L2ers.

The mixed-effects model for the late L2ers revealed a significant effect of *Construction* ($\beta = -5.051$, $SE = 1.374$, $p < .001$) and *Interpretation* ($\beta = 5.205$, $SE = 1.500$, $p = .001$), as shown in Table 6.13: Gapping conditions had higher rates of acceptance than VPE conditions and SR conditions had higher rates of acceptance than the OR conditions.

Table 6.13
Results of the Logistic Mixed-Effects Regression for the Late L2er Judgment Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.714	0.791	.030
Construction	−5.051	1.374	< .001
Interpretation	5.205	1.500	.001
Construction × Interpretation	18.662	2.590	< .001

Note. Model formula: glmer(Judgment ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

More importantly for the current study, a significant interaction was found between *Construction* and *Interpretation* ($\beta = 18.662$, $SE = 2.590$, $p < .001$). To unpack this interaction, I constructed follow-up mixed-effects models for pairwise comparisons. These models found a significantly lower acceptance rate for the VPE-OR condition relative to both the VPE-SR condition ($\beta = 10.888$, $SE = 4.080$, $p = .008$) and the Gapping-OR condition ($\beta = -41.709$, $SE = 10.334$, $p < .001$). These results indicate that late L2ers were able to reject VPE sentences with an object reading. The Gapping-SR condition also had a lower acceptance rate than the Gapping-OR condition ($\beta = -11.488$, $SE = 5.999$, $p = .056$), but this difference was only marginally significant. There was no significant difference between the VPE-SR condition and the Gapping-SR condition ($\beta = 4.087$, $SE = 3.102$, $p = .188$).

The late L2ers, by contrast, displayed no effect of *Proficiency* ($\beta = -0.236$, $SE = 0.339$, $p = .487$) in the model constructed with the *Proficiency* factor added, as shown in Table 6.14.

Table 6.14
Results of the Logistic Mixed-Effects Regression for the Late L2er Judgment Data in Study 3 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.445	0.708	.001
Construction	−2.298	1.282	.073
Interpretation	3.847	1.330	.004
Proficiency	−0.236	0.339	.487
Construction × Interpretation	13.324	2.471	< .001
Construction × Proficiency	−0.106	0.653	.871
Interpretation × Proficiency	−0.265	0.672	.693
Construction × Interpretation × Proficiency	−0.360	1.253	.774

Note. Model formula: glmer(Judgment ~ Construction * Interpretation * Proficiency + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Nor were there significant interactions between *Construction* and *Proficiency* ($\beta = -0.106$, $SE = 0.653$, $p = .871$), between *Interpretation* and *Proficiency* ($\beta = -0.265$, $SE = 0.672$, $p = .693$), or among *Construction*, *Interpretation*, and *Proficiency* ($\beta = -0.360$, $SE = 1.253$, $p = .774$).

Despite the fact that the mixed-effects model for the late L2ers did not yield a significant effect of *Proficiency*, I ran a series of simple regression analyses to investigate whether there might be a relationship between their proficiency scores and their sensitivity scores (for the individual sensitivity scores, see Appendix J) and/or their acceptance rates in each of the four conditions. These analyses did not reveal a significant effect of *Proficiency* on any of the following: the sensitivity scores ($\beta = -1.499$, $SE = 1.880$, $p = .432$), the acceptance rate for the VPE-SR condition ($\beta = -0.537$, $SE = 0.674$, $p = .432$), the acceptance rate for the VPE-OR condition ($\beta = -0.008$, $SE = 2.959$, $p = .486$), the acceptance rate for the Gapping-SR condition ($\beta = -2.087$, $SE = 3.114$, $p = .488$), or the acceptance rate for the Gapping-OR condition ($\beta = -1.850$, $SE = 2.705$, $p = .500$).

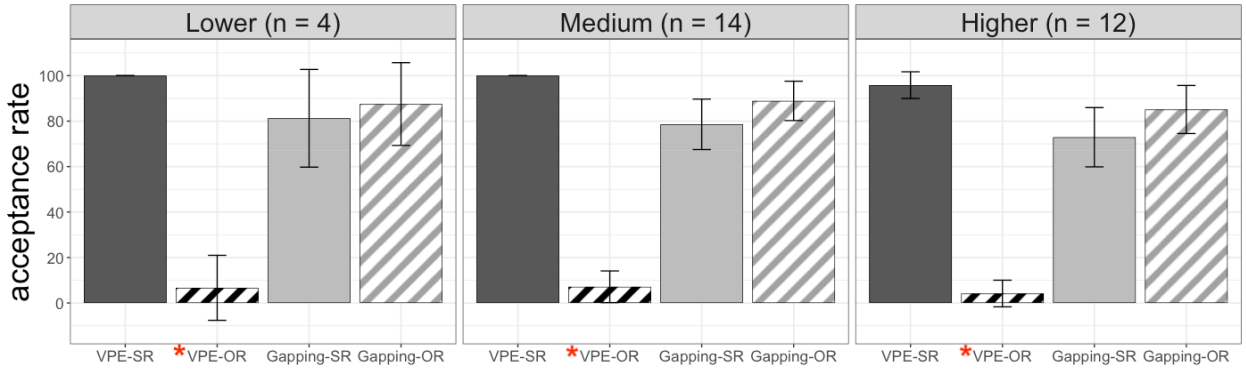
The late L2ers were also divided into three proficiency groups for an additional round of analysis, as shown in Table 6.15 and Figure 6.8. There were no pronounced differences between the groups.

Table 6.15
Late L2ers' Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Proficiency Group

	VPE-SR	***VPE-OR**	**Gapping-SR**	**Gapping-OR**
Lower	100.00	6.67	81.25	87.50
(*n* = 4)	(*SD* = 0.00)	(*SD* = 25.82)	(*SD* = 40.31)	(*SD* = 34.16)
Medium	100.00	7.14	78.57	88.89
(*n* = 14)	(*SD* = 0.00)	(*SD* = 25.99)	(*SD* = 41.40)	(*SD* = 31.72)
Higher	95.83	4.17	72.92	85.11
(*n* = 12)	(*SD* = 20.19)	(*SD* = 20.19)	(*SD* = 44.91)	(*SD* = 35.99)

Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading.

Figure 6.8
Late L2ers' Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Proficiency Group



Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading. Error bars represent 95% confidence intervals.

In sum, late L2ers, regardless of proficiency, accepted VPE with a subject reading, Gapping with a subject reading, and Gapping with an object reading but rejected (impossible) VPE with an object reading.

6.4 Discussion

This study investigated the acquisition of the contrast between possible vs. impossible interpretations of VPE and Gapping by L1-English children and L1-Korean L2ers of English. The results of the picture-sentence matching task showed that L1 children as young as 6 come to display the target contrast and that early L2ers with higher proficiency and most late L2ers are capable of developing a sensitivity to the contrast, too.

The remainder of this section is organized as follows. First, I discuss three possible explanations for the finding that each of the four groups had different rates of acceptance for the Gapping-SR condition and the Gapping-OR condition even though both conditions are grammatical. I then propose a sequence of developmental stages for the interpretation contrast between VPE and Gapping based on my analysis of individual response data from the L1 children and the early and late L2ers.

6.4.1 Processing difficulty of Gapping with a subject reading.

One noteworthy finding of this study is that for Gapping sentences, all four of the groups accepted the subject reading less often than the object reading. It might be tempting to conclude that this result is in some way attributable to grammar, but I do not believe that this is the case. Even though the Gapping-SR condition was accepted at lower rates than the Gapping-OR condition, it was still accepted significantly more often than the ungrammatical VPE-OR condition (L1 adults: $\beta = -8.701$, $SE = 2.801$, $p = .002$; L1 children: $\beta = -3.205$, $SE = 0.851$, $p < .001$; early L2ers: $\beta = -2.336$, $SE = 0.615$, $p = .001$; late L2ers: $\beta = -6.860$, $SE = 2.428$, $p < .001$). It is therefore reasonable to consider alternative explanations for this pattern of results. I discuss three below.

The first explanation I would like to consider is related to information processing, more specifically, processing in relation to information structure. Kobayashi (2005) states that parsers prefer to resolve the contrastive remnant (the element stranded after Gapping has applied; e.g., *Dad* in *Mom hugged the boy at home and Dad in the park*.), which is new information, as an object rather than as a subject because an object position often serves as the position for new information (see also Harris & Carlson, 2018).

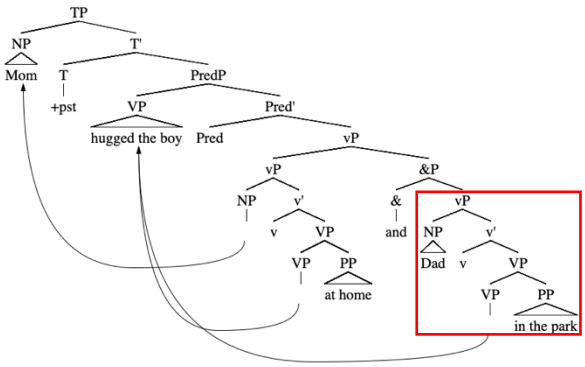
Another possible explanation concerns topic continuity. According to William O'Grady (personal communication, 11 October 2019), "in the absence of evidence to the contrary, the processor prefers coordinate structures to have identical subjects (since this favors topic continuity)." It is possible that parsers want to retain the subject already processed in the first conjunct, and thus they resolve the first argument in the gapped conjunct as an object rather than as a subject.

Lastly, the lower rates of acceptance for the Gapping-SR condition relative to the Gapping-OR condition are consistent with a Minimal Attachment account (Frazier, 1978, 1987;

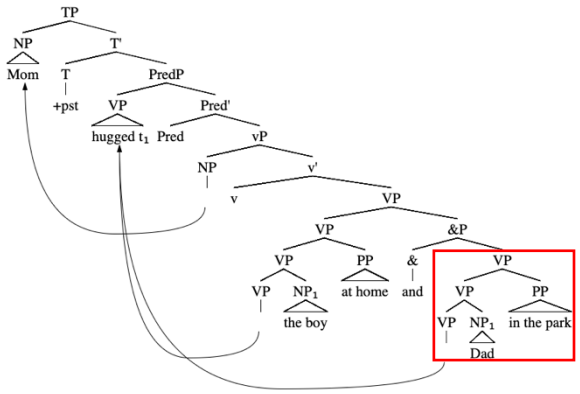
Gibson, 1998) according to which parsers prefer the simplest syntactic analysis consistent with the word string. For example, when parsers encounter a sentence such as *Mom hugged the boy at home and Dad in the park*, they prefer to keep the second conjunct (i.e., *Dad in the park*) as small as possible. As shown in (2a) and (2b), the conjunct is structurally smaller in the Gapping-OR sentence than in its Gapping-SR counterpart.

(2) Analysis of Gapping (adapted from the analyses of Johnson, 2000, 2009 & Zoerner, 1999)

a. Gapping-SR



b. Gapping-OR



What the three explanations above have in common is that they view the preference of the object reading over the subject reading for Gapping in terms of processing efficiency. To ascertain whether processing pressures may have contributed to the pattern of results observed in the current study, I analyzed the RT data for each group of participants. As mentioned in §5.4.1, I follow the existing psycholinguistic literature in assuming that longer RTs are a sign of processing challenges associated with dependent propositional relations, complex syntactic structures, encoding/retrieval interferences, and so on (for an eye-tracking-while-reading paradigm, see Kliegl & Laubrock, 2018; for a grammaticality judgment paradigm, see Smith, 2011; for a self-paced listening picture-verification paradigm, see Peristeri & Tsimpli, 2013; for a self-paced reading paradigm, see Villata et al., 2018).

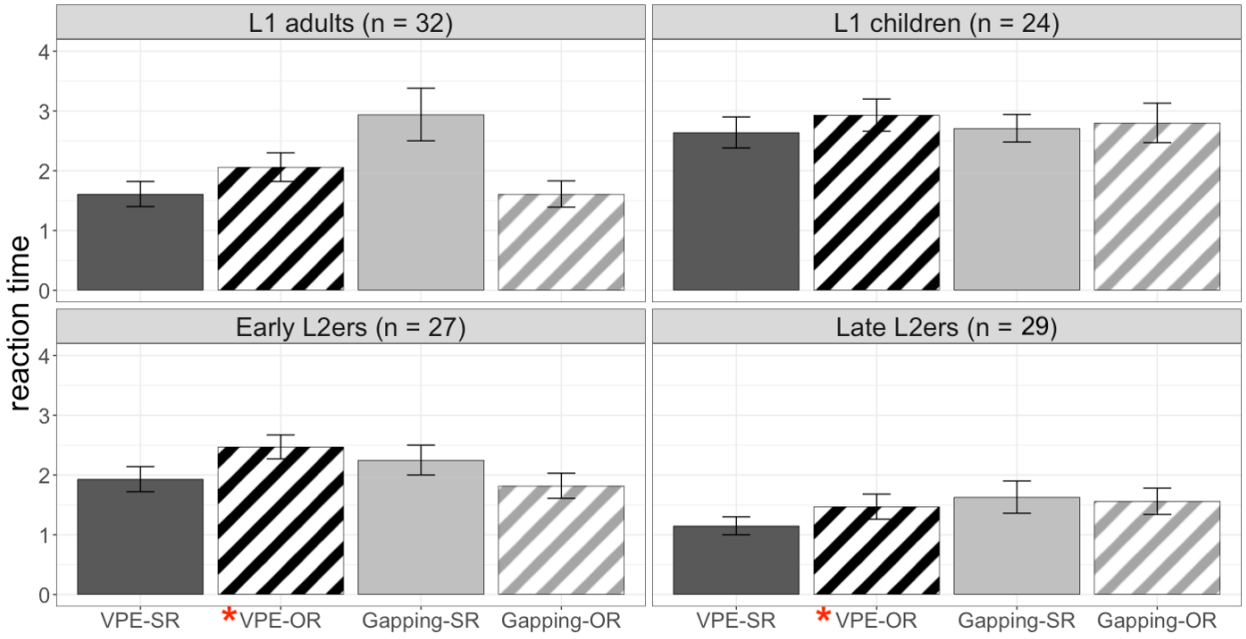
Table 6.16 and Figure 6.9 provide a general picture of the RT patterns for each group.

Table 6.16
Mean Reaction Times (in Seconds) in the Picture-Sentence Matching Task per Condition and Group

	VPE-SR	*VPE-OR	Gapping-SR	Gapping-OR
L1 Adults	1.61	2.06	2.94	1.61
(*n* = 32)	(*SD* = 1.18)	(*SD* = 1.29)	(*SD* = 1.99)	(*SD* = 1.19)
L1 Children	2.64	2.93	2.71	2.80
(*n* = 24)	(*SD* = 1.17)	(*SD* = 1.22)	(*SD* = 1.02)	(*SD* = 1.49)
Early L2ers	1.93	2.47	2.25	1.82
(*n* = 27)	(*SD* = 1.02)	(*SD* = 1.00)	(*SD* = 1.15)	(*SD* = 0.99)
Late L2ers	1.15	1.47	1.63	1.56
(*n* = 29)	(*SD* = 0.77)	(*SD* = 1.07)	(*SD* = 1.24)	(*SD* = 1.08)

Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading.

Figure 6.9
Mean Reaction Times (in Seconds) in the Picture-Sentence Matching Task per Condition and Group



Notes. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading. Error bars represent 95% confidence intervals.

The shortest RTs for the adults were in the VPE-SR condition and the Gapping-OR condition. For the L1 children and the late L2ers, the shortest RTs were in the VPE-SR condition. It was the Gapping-OR condition that had the shortest RTs in the early L2ers. These results indicate that VPE sentences with a subject reading and/or Gapping sentences with an object reading were

unproblematic for these groups to process and hence unproblematic for them to judge relative to the events depicted in the picture. On the other hand, the L1 children and early L2ers had their longest RTs in the impossible VPE-OR condition. But the L1 adults and late L2ers, by contrast, both had their longest RTs in the Gapping-SR condition, presumably due to the processing pressures discussed above.

The RT data for each group were further examined using linear mixed-effects regression. Henceforth, I only discuss effects that reached at least marginal significance, but all non-significant values can be found in the results table. The model for the L1 adults, provided in Table 6.17, revealed a marginally significant effect of *Construction* ($\beta = -0.530$, $SE = 0.256$, $p = .055$) and *Interpretation* ($\beta = 0.481$, $SE = 0.244$, $p = .066$), with longer RTs in the Gapping conditions than in the VPE conditions and with longer RTs in the SR conditions than in the OR conditions.

Table 6.17
Results of the Linear Mixed-Effects Regression for the L1 Adult Reaction Time Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.114	0.131	< .001
Construction	−0.530	0.256	.055
Interpretation	0.481	0.244	.066
Construction × Interpretation	−2.056	0.400	< .001

Note. Model formula: lmer(RT ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Furthermore, there was a significant interaction between *Construction* and *Interpretation* ($\beta = -2.056$, $SE = 0.400$, $p < .001$). Follow-up mixed-effects models showed significantly longer RTs in the Gapping-SR condition than both the VPE-SR condition ($\beta = -1.646$, $SE = 0.404$, $p < .001$) and the Gapping-OR condition ($\beta = 1.623$, $SE = 0.423$, $p = .001$). In addition, RTs in the VPE-OR condition were significantly longer than those in the VPE-SR condition ($\beta = -0.459$, $SE = 0.180$, $p = .024$) and those in the Gapping-OR condition ($\beta = 0.448$, $SE = 0.157$, $p = .005$).

Interestingly, the L1 adults had longer RTs in the possible Gapping-SR condition and the impossible VPE-OR condition than in the other two conditions. The longer RTs in the Gapping-SR condition relative to the VPE-SR and Gapping-OR conditions are likely the result of increased processing load associated with dispreferred information structure, topic

discontinuity, or structural complexity (see above). By contrast, the longer RTs in the VPE-OR condition relative to the VPE-SR and Gapping-OR conditions may have resulted from a different type of processing difficulty, this time associated with online and/or post-interpretive processes (e.g., Caplan & Waters, 1999). Recall that it was only in the (impossible) VPE-OR condition that the interpretations that participants got from the target sentence and the context sentences (and pictures) did not match. Because generating and comparing different interpretations is not an instantaneous process, it might be the case that L1 adults simply needed more time to figure out that the interpretation that they assigned to the target sentence (i.e., the subject reading) was not compatible with the interpretation of the context sentences (and pictures).[2]

The mixed-effects model for the L1 children did not reveal any significant effects, as shown in Table 6.18.

Table 6.18
Results of the Linear Mixed-Effects Regression for the L1 Child Reaction Time Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.849	0.146	< .001
Construction	−0.005	0.225	0.984
Interpretation	−0.202	0.135	0.171
Construction × Interpretation	−0.176	0.396	0.660

Note. Model formula: lmer(RT ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

A separate linear mixed-effects model with the factor *Age* added also did not show any significant effect related to *Age*, as shown in Table 6.19.

[2] My thanks to Bonnie D. Schwartz (personal communication, 9 April 2020) for discussing this issue with me.

Table 6.19
Results of the Linear Mixed-Effects Regression for the L1 Child Reaction Time Data in Study 3 with the Factor Age *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	4.248	1.704	.020
Construction	−0.343	1.847	.854
Interpretation	1.038	1.930	.597
Age	−0.216	0.261	.416
Construction × Interpretation	4.043	4.880	.416
Construction × Age	0.052	0.281	.854
Interpretation × Age	−0.191	0.295	.525
Construction × Interpretation × Age	−0.648	0.746	.394

Note. Model formula: lmer(RT ~ Construction * Interpretation * Age + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Regarding the RT data from the early L2ers, there was a significant interaction between *Construction* and *Interpretation* ($\beta = -0.880$, $SE = 0.273$, $p = .005$), as shown in Table 6.20.

Table 6.20
Results of the Linear Mixed-Effects Regression for the Early L2er Reaction Time Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.167	0.119	< .001
Construction	0.141	0.110	.212
Interpretation	−0.050	0.117	.672
Construction × Interpretation	−0.880	0.273	.005

Note. Model formula: lmer(RT ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Follow-up mixed effects models revealed that there were significantly longer RTs in the VPE-OR condition than in either the VPE-SR condition ($\beta = -0.480$, $SE = 0.147$, $p = .004$) or the Gapping-OR condition ($\beta = 0.631$, $SE = 0.161$, $p < .001$), thus indicating that early L2ers experienced processing difficulty when making judgments on the interpretation of VPE-OR sentences relative to the contexts (and pictures). As was the case for the L1 adults, it is likely that the early L2ers also needed more time to figure out that their interpretation of the target sentence did not match the interpretation they had constructed for the context sentences (and pictures). Furthermore, the RTs in the Gapping-SR condition were significantly longer than those in both the VPE-SR condition ($\beta = -0.304$, $SE = 0.140$, $p = .041$) and the Gapping-OR condition

(β = 0.434, *SE* = 0.165, p = .014). This result suggests that the early L2ers experienced processing pressures like the L1 adults did, as discussed above.

The mixed-effects model which added the factor *Proficiency* found a significant interaction among *Construction*, *Interpretation*, and *Proficiency* (β = −0.177, *SE* = 0.083, p = .043) (see Table 6.21).

Table 6.21
Results of the Linear Mixed-Effects Regression for the Early L2er Reaction Time Data in Study 3 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	2.130	0.123	< .001
Construction	0.216	0.113	.072
Interpretation	−0.064	0.124	.611
Proficiency	−0.043	0.042	.313
Construction × Interpretation	−1.045	0.269	.002
Construction × Proficiency	0.067	0.042	.124
Interpretation × Proficiency	−0.049	0.047	.311
Construction × Interpretation × Proficiency	−0.177	0.083	.043

Note. Model formula: lmer(RT ~ Construction * Interpretation * Proficiency + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Further simple regression analyses revealed that this three-way interaction mainly came from higher-proficiency L2ers' significantly faster RTs for the VPE-SR condition (β = −0.111, *SE* = 0.050, p = .037). However, the analyses found no significant effect of *Age* on the RTs for the VPE-OR condition (β = 0.056, *SE* = 0.058, p = .352), the Gapping-SR condition (β = −0.057, *SE* = 0.069, p = .410), or the Gapping-OR condition (β = −0.095, *SE* = 0.065, p = .158).

The results from the mixed-effects model for the late L2ers, provided in Table 6.22, did not show any significant effect of *Construction* or *Interpretation*.

Table 6.22
Results of the Linear Mixed-Effects Regression for the Late L2er Reaction Time Data in Study 3

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.485	0.096	< .001
Construction	−0.297	0.218	.191
Interpretation	−0.193	0.170	.272
Construction × Interpretation	−0.389	0.367	.303

Note. Model formula: lmer(RT ~ Construction * Interpretation + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

Furthermore, no significant effect of *Proficiency* emerged in the separate model with the *Proficiency* factor added (see Table 6.23).

Table 6.23
Results of the Linear Mixed-Effects Regression for the Late L2er Reaction Time Data in Study 3 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	1.477	0.105	< .001
Construction	−0.274	0.222	.234
Interpretation	−0.242	0.176	.187
Proficiency	0.011	0.050	.829
Construction × Interpretation	−0.480	0.382	.223
Construction × Proficiency	−0.029	0.060	.636
Interpretation × Proficiency	0.066	0.060	.281
Construction × Interpretation × Proficiency	0.119	0.131	.375

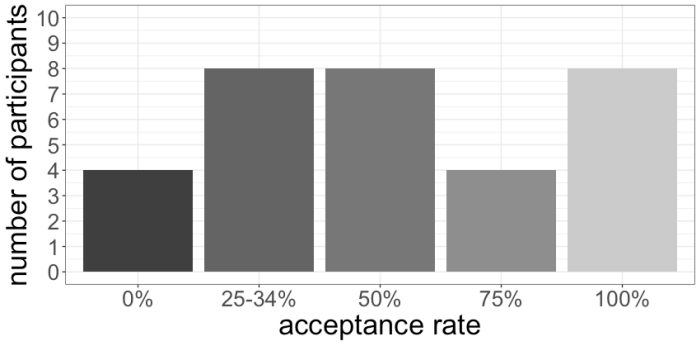
Note. Model formula: lmer(RT ~ Construction * Interpretation * Proficiency + (1 + Construction * Interpretation | participant) + (1 + Construction * Interpretation | item))

One important finding of the by-group RT analyses was that the L1 adults and the early L2ers had significantly longer RTs in the Gapping-SR condition than in the other conditions. There was no statistically significant pattern found in the RT data from the L1 children or the late L2ers.

Crucially, the finding that the L1 adults had both longer RTs in the Gapping-SR condition than the other conditions and higher acceptance rates for the Gapping-SR condition than the VPE-OR condition indicates that their lower rate of acceptance in this condition results from processing rather than grammar. Additional evidence for this processing account comes from the distribution of acceptance rates from individual L1 adults given in Figure 6.10, which

shows that the majority in that group (62.5%) accepted the Gapping-SR condition 50% of the time or more. For this processing-related reason, then, the acceptance rates for the Gapping-SR condition may not provide a clear picture of the participants' grammatical knowledge regarding Gapping with a subject reading.

Figure 6.10
Distribution of L1 Adult Acceptance Rates for Gapping-SR



6.4.2 Developmental sequences for the interpretation contrast between VP-Ellipsis and Gapping.

This section proposes possible developmental sequences for the interpretation contrast between VPE and Gapping in each learner group by examining the response patterns of individual participants. As was done previously (see §5.4.2), *Age* served as a guideline for the L1 children and *Proficiency* served as a guideline for the early and late L2ers. Given the variety of response patterns for the Gapping-SR sentences (arguably due to processing difficulty—see discussion in §6.4.1), including the Gapping-SR condition in the individual analysis would make it difficult to interpret the results. For this reason, I analyzed only the data from the other three conditions (i.e., VPE-SR, VPE-OR, and Gapping-OR).

The four criteria that I used for this analysis are presented in Table 6.24.

Table 6.24
Criteria for Individual Analysis in Study 3

Criterion	Item
A	Knowledge of the grammaticality of VPE with a subject reading
B	Knowledge of the grammaticality of Gapping with an object reading
C	Knowledge of the contrast between VPE with a subject reading and *VPE with an object reading
D	Knowledge of the ungrammaticality of *VPE with an object reading

For Criteria A, B, and D, I used the comparison-against-chance diagnostic based on a binomial distribution (see §5.4.2). There were four items per condition in the picture-sentence matching task. The probability of having a correct response was 0.33 since there are three response options (i.e., smiling face, frowning face, and question mark). It turns out that all four items need to be judged correctly ($p = .012$)[3] for accuracy to exceed chance performance. However, I decided to set the criterion at *having at least 3 out of 4 items correct*—which still reaches marginal significance ($p = 0.099$)[4]—because there were seven different L1 adults (out of 32) who responded correctly to only three out of four items in one or two conditions (two for the VPE-SR condition, three for the VPE-OR condition, and four for the Gapping-OR condition). For Criterion C, I compared the numbers of VPE-SR and VPE-OR items that the participant judged correctly; in order to qualify as "pass," the number of correct items for (possible) VPE-SR had to exceed that for (impossible) VPE-OR.

Participants received a check mark (✓) for each criterion they passed. I identified developmental stages based on the following requirements: (a) Each developmental stage needed to start with at least two participants of consecutive age/proficiency scores who passed the criterion under inspection and (b) there could not be more than two consecutive participants within a developmental stage who failed to pass the criterion. The developmental stages identified for the L1 children, the early L2ers, and the late L2ers are shown in, respectively, Tables 6.25–6.27.

[3] The *p*-value calculated based on a binomial distribution is: {4!/(4!*0!)}*0.33^4*(1–0.33)^0 = 0.012.
[4] The *p*-value calculated based on a binomial distribution is: {4!/(3!*1!)}*0.33^3*(1–0.33)^1 = 0.099.

Table 6.25

*Developmental Sequence of the Interpretation Contrast between VPE and Gapping: L1 Children in Chronological Order (*n = *24)*

Age	Participant	Knowledge of the possibility of VPE-SR	Knowledge of the possibility of Gapping-OR	Knowledge of the contrast between VPE-SR and *VPE-OR	Knowledge of the impossibility of *VPE-OR	Developmental stage
5;4	L1C_10	✓	✓			(1) Knowledge of the possibility of VPE-SR and Gapping-OR
5;6	L1C_28	✓	✓	✓	✓	(2) Knowledge of the contrast between VPE-SR and *VPE-OR
5;8	L1C_25	✓	✓	✓		
5;10	L1C_13	✓		✓	✓	(3) Knowledge of the impossibility of *VPE-OR
5;11	L1C_14		✓	✓	✓	
5;11	L1C_26	✓	✓	✓		
6;1	L1C_20			✓	✓	
6;3	L1C_17	✓	✓	✓		
6;3	L1C_24	✓	✓	✓	✓	
6;6	L1C_31	✓	✓	✓		
6;6	L1C_15	✓	✓	✓	✓	
6;7	L1C_18	✓	✓			
6;8	L1C_12	✓	✓	✓	✓	
6;8	L1C_27	✓	✓	✓	✓	
6;9	L1C_33	✓	✓	✓	✓	
6;9	L1C_23	✓	✓	✓	✓	
6;10	L1C_16	✓	✓	✓	✓	
6;10	L1C_30	✓	✓	✓	✓	
6;11	L1C_11		✓			
6;11	L1C_21	✓	✓	✓	✓	
7;1	L1C_22	✓	✓	✓	✓	
7;1	L1C_32	✓	✓	✓	✓	
7;2	L1C_29	✓	✓	✓	✓	
7;2	L1C_19		✓	✓	✓	

Notes. The check sign (✓) indicates presence of knowledge. The four criteria were as follows: knowledge of the possibility of VPE with a subject reading (VPE-SR; 3–4 correct out of 4); knowledge of the possibility of Gapping with an object reading (Gapping-OR; 3–4 correct out of 4); knowledge of the contrast between VPE-SR and *VPE-OR (correct VPE-SR > correct *VPE-OR); knowledge of the impossibility of VPE with an object reading (*VPE-OR; 3–4 correct out of 4).

Table 6.26

*Developmental Sequence of the Interpretation Contrast between VPE and Gapping: Early L2ers in Proficiency Order (*n *= 27)*

Proficiency	Participant	Knowledge of the possibility of VPE-SR	Knowledge of the possibility of Gapping-OR	Knowledge of the contrast between VPE-SR and *VPE-OR	Knowledge of the impossibility of *VPE-OR	Developmental stage
−6.65 (L)	EL2_11	✓	✓			(1) Knowledge of the possibility of VPE-SR and Gapping-OR
−5.98 (L)	EL2_04	✓	✓			
−3.80 (L)	EL2_23	✓	✓	✓		
−3.41 (L)	EL2_12				✓	
−3.25 (L)	EL2_24	✓				
−2.65 (L)	EL2_05	✓	✓	✓		(2) Knowledge of the contrast between VPE-SR and *VPE-OR
−2.52 (L)	EL2_21	✓	✓	✓	✓	
−2.49 (L)	EL2_10		✓			
−2.36 (L)	EL2_07				✓	
−2.35 (L)	EL2_08	✓	✓	✓		
−1.96 (L)	EL2_27	✓	✓	✓	✓	
−1.10 (M)	EL2_18	✓	✓			
−0.78 (M)	EL2_25	✓	✓	✓		
−0.71 (M)	EL2_15	✓	✓	✓	✓	
−0.19 (M)	EL2_02	✓	✓	✓		
0.39 (M)	EL2_03		✓			
0.51 (M)	EL2_06	✓	✓	✓	✓	(3) Knowledge of the impossibility of *VPE-OR
0.51 (M)	EL2_26	✓	✓	✓	✓	
0.53 (M)	EL2_01	✓	✓	✓		
0.80 (M)	EL2_22	✓	✓	✓	✓	
1.45 (H)	EL2_09	✓	✓	✓	✓	
1.62 (H)	EL2_13	✓	✓	✓	✓	
1.69 (H)	EL2_19	✓	✓	✓		
1.90 (H)	EL2_17	✓	✓	✓	✓	
2.79 (H)	EL2_20	✓	✓	✓	✓	
2.90 (H)	EL2_16	✓	✓	✓	✓	
3.51 (H)	EL2_14	✓	✓	✓	✓	

Notes. (L), (M), and (H) respectively indicate the Lower proficiency group, the Medium proficiency group, and the Higher proficiency group. The check sign (✓) indicates presence of knowledge. The four criteria were as follows: knowledge of the possibility of VPE with a subject reading (VPE-SR; 3–4 correct out of 4); knowledge of the possibility of Gapping with an object reading (Gapping-OR; 3–4 correct out of 4); knowledge of the contrast between VPE-SR and *VPE-OR (correct VPE-SR > correct *VPE-OR); knowledge of the impossibility of VPE with an object reading (*VPE-OR; 3–4 correct out of 4).

Table 6.27

*Developmental Sequence of the Interpretation Contrast between VPE and Gapping: Late L2ers in Proficiency Order (*n = *30)*

Proficiency	Participant	Knowledge of the possibility of VPE-SR	Knowledge of the possibility of Gapping-OR	Knowledge of the contrast between VPE-SR and *VPE-OR	Knowledge of the impossibility of *VPE-OR	Developmental stage
−4.36 (L)	LL2_26	✓	✓	✓	✓	N/A
−2.31 (L)	LL2_22	✓	✓	✓	✓	
−2.00 (L)	LL2_05	✓	✓	✓	✓	
−1.36 (L)	LL2_24	✓	✓	✓	✓	
−0.79 (M)	LL2_18	✓	✓	✓	✓	
−0.31 (M)	LL2_08	✓	✓	✓	✓	
−0.24 (M)	LL2_19	✓	✓	✓	✓	
0.14 (M)	LL2_01	✓	✓	✓		
0.14 (M)	LL2_20	✓	✓	✓	✓	
0.22 (M)	LL2_16	✓	✓	✓	✓	
0.24 (M)	LL2_14	✓	✓	✓	✓	
0.25 (M)	LL2_04	✓		✓	✓	
0.46 (M)	LL2_11	✓	✓	✓	✓	
0.59 (M)	LL2_21	✓	✓	✓	✓	
0.64 (M)	LL2_10	✓	✓	✓	✓	
0.80 (M)	LL2_27	✓	✓	✓	✓	
0.72 (M)	LL2_28	✓	✓	✓	✓	
0.90 (M)	LL2_02	✓	✓	✓	✓	
1.44 (H)	LL2_23	✓	✓	✓	✓	
1.44 (H)	LL2_06	✓	✓	✓	✓	
1.52 (H)	LL2_13	✓	✓	✓	✓	
1.85 (H)	LL2_17	✓	✓	✓	✓	
2.09 (H)	LL2_07	✓	✓	✓	✓	
2.16 (H)	LL2_15	✓	✓	✓	✓	
2.39 (H)	LL2_09	✓	✓	✓	✓	
2.39 (H)	LL2_29	✓	✓	✓	✓	
2.41 (H)	LL2_25	✓		✓	✓	
2.80 (H)	LL2_03	✓	✓	✓	✓	
3.53 (H)	LL2_12	✓		✓	✓	
3.84 (H)	LL2_30	✓	✓	✓	✓	

Notes. (L), (M), and (H) respectively indicate the Lower proficiency group, the Medium proficiency group, and the Higher proficiency group. The check sign (✓) indicates presence of knowledge. The four criteria were as follows: knowledge of the possibility of VPE with a subject reading (VPE-SR; 3–4 correct out of 4); knowledge of the possibility of Gapping with an object reading (Gapping-OR; 3–4 correct out of 4); knowledge of the contrast between VPE-SR and *VPE-OR (correct VPE-SR > correct *VPE-OR); knowledge of the impossibility of VPE with an object reading (*VPE-OR; 3–4 correct out of 4).

Three distinct developmental stages were identified for the L1 children. In the first stage, the onset of which was age 5;4, the children showed they knew that both VPE-SR and Gapping-OR are possible. Next came knowledge of the interpretation contrast between VPE-SR

and VPE-OR at age 5;6. In the last stage, which began at age 5;10, they know that the object reading of VPE is not possible.

The same three stages were observed in the data from the early L2ers. First, they acquired knowledge that VPE-SR and Gapping-OR are possible. Then they differentiated (possible) VPE-SR from (impossible) VPE-OR interpretations. Finally, they exhibited knowledge of the impossibility of VPE-OR.

Perhaps because of their high proficiency scores, the majority of the late L2ers passed all four of the criteria. It was therefore not possible to identify distinct developmental stages in the data from these participants. This pattern of results is similar to the one I observed for the same group of participants in Study 2, the acceptability judgment study (see §5.4.2).

6.5 Conclusion

To summarize Study 2: The older L1 children and the higher-proficiency L2ers were shown to have acquired the interpretation contrast between VPE and Gapping in English. Analysis of the data from individual participants revealed that 13 out of 24 L1 children had the target contrast, and the youngest child among those was aged 5;6. Regarding the early L2ers, a *Proficiency* effect was observed in which the ungrammatical VPE-OR condition was rejected more often by higher-proficiency L2ers. The individual data analysis showed that 12 out of 27 early L2ers and 26 out of 30 late L2ers showed clear evidence of having acquired the target contrast. These results indicate that children and (early and late) L1-Korean L2ers can come to have implicit knowledge of the interpretation contrast between VPE and Gapping despite the learnability challenges they face.

CHAPTER VII
STUDY 4: L2 PROCESSING OF GAPPING

This chapter reports on a processing study which investigated whether L1-Korean adult L2ers of English postulate a verb gap and reconstruct verb information at the gap site as they process Gapping in real time.[1] I first present the research questions of this study in §7.1. Then I lay out a method for addressing them in §7.2 and spell out my predictions in §7.3. Section 7.4 illustrates the data analysis method, and §7.5 and §7.6, respectively, report and discuss the results. Section 7.7 concludes this chapter.

7.1 Research Questions

I ask the following research questions in this study:

> Can adult L1-Korean L2ers of English come to recognize and resolve a verb gap in Gapping sentences in real time? What role does L2 proficiency play?

To answer these questions, I conducted a self-paced reading task in a plausibility manipulation paradigm (Garnsey et al., 1989). I also implemented an independent proficiency test in order to investigate an effect of L2 proficiency in the L2ers' processing patterns.

7.2 Method

7.2.1 Participants.

Two groups of participants took part in this study online through Ibex Farm (Drummond, 2007): Fifty-nine L1-Korean L2ers of English (L2-English group), all of them recruited either at universities in Korea or at the University of Hawai‘i at Mānoa, and 65 English native speakers as a control group (L1-English group), all recruited from the University of Hawai‘i at Mānoa. None of them participated in Study 2 (Chapter 5) or Study 3 (Chapter 6). Eleven of the L2ers and 12 of the native speakers were excluded in the statistical analyses because their accuracy rate for the comprehension questions in the self-paced reading task was lower than 80%. The mean age of

[1] The pilot version of this study was conducted together with Amber Camp (University of Hawai‘i at Mānoa) in Fall 2018.

English onset for the remaining 48 L2ers was 7.85 (*SD* = 1.96). Their mean length of residence in English-speaking countries was 10.31 months (*SD* = 20.42), and their mean score from an English cloze test (J. D. Brown, 1980; see Appendix K), the proficiency measure used in this study, was 30.71 (*SD* = 8.84) out of 50 (which indicates relatively high proficiency—cf. the native speakers' mean of 38.89 (*SD* = 5.88)). The English curriculum in Korea (Ministry of Education, 2015) is implemented to students from the 3rd grade to the 12th grade. The goal of this curriculum is to develop students' communicative competence in listening, speaking, reading, and writing. The curriculum provides 2 lesson hours (80 minutes) per week to 3rd and 4th graders, 3 lesson hours (120 minutes) per week to 5th and 6th graders, approximately 3–4 lesson hours (135–180 minutes) per week to 7th to 9th graders, and 4.5–5 lesson hours (225–250 minutes) per week to 10th to 12th graders. Participant background information for each group is given in Table 7.1.

Table 7.1
Background Information for Participants in Study 4

	Age at Testing	**Age of English Onset**	**Length of residence in an English-speaking country in months**	**Cloze test score**
English native speakers (*n* = 53)	21.58 (*SD* = 4.72; range = 18–46)	N/A	N/A	38.89 (*SD* = 5.88; range = 22–50)
L1-Korean L2ers of English (*n* = 48)	22.71 (*SD* = 2.95; range = 18–29)	7.85 (*SD* = 1.96; range = 4–13)	10.31 (SD = 20.42; range = 0–96)	30.71 (*SD* = 8.84; range = 11–47)

7.2.2 Materials.

For this experiment, I crossed *Construction* (Gapping; VP-Ellipsis (VPE)) and *Plausibility* of the verb for the NP object in the gapped clause (Plausible; Implausible) in a 2 × 2 Latin square design, as illustrated in Table 7.2. This resulted in four experimental lists, each containing 20 critical sentences and 50 fillers. The 20 critical sentences were constructed based on the ones used in Kaan et al.'s (2004) study. A full list of all experimental sentences is in Appendix L.

Table 7.2
Segmentation of a Sample Sentence for Each Condition

Condition \ Region	1	2	3	4	5	6	7	8 critical region	9 spill-over region	10
(a) Gapping-P	Bill	**ordered**	coffee	and	tea	at the cafe,	and	Jane [e] sandwiches	and cake	at the bakery.
(b) *Gapping-I	* Bill	**drank**	coffee	and	tea	at the cafe,	and	Jane [e] sandwiches	and cake	at the bakery.
(c) VPE-P (baseline)	Bill	**ordered**	coffee	and	tea	at the cafe,	and	Jane did [e]	too	with his brother.
(d) VPE-I (baseline)	Bill	**drank**	coffee	and	tea	at the cafe,	and	Jane did [e]	too	with his brother.

Notes. The VPE-I condition sentences were not themselves implausible but instead provided a baseline against which the *Gapping-I condition sentences could be compared.
Gapping-P = Gapping-Plausible, *Gapping-I = Gapping-Implausible; VPE-P = VPE-Plausible, VPE-I = VPE-Implausible.

In order to stretch out the verb-gap processing region in the Gapping conditions, the direct objects in the gapped clause (as well as the non-gapped clause) were all conjoined NPs in the form of *NP and NP*. Accordingly, Segment 8 and Segment 9 constituted my regions of interest. The former was the critical region, where participants might first recognize a verb gap and attempt to resolve the verb and gap dependency. The latter served as a spill-over region, allowing us to see any delayed effects of gap processing for L2ers, who have been found to process sentences at a slower rate than native speakers (e.g., Dallas, 2008; Dekydtspotter, Schwartz, & Sprouse, 2006; Felser, Cunnings, Batterham, & Clahsen, 2012; Hopp, 2009, 2010). At one or both of these regions, I predicted that there would be a plausibility effect between only the two Gapping conditions, specifically a reading-time slowdown in the Gapping-Implausible condition compared to the Gapping-Plausible condition.

The two VPE conditions were included as the baseline to be contrasted with the Gapping conditions. Note that the VPE-Implausible condition sentences were not themselves implausible but instead provided a baseline against which the Gapping-Implausible condition sentences could be compared. VPE was selected as a control construction for the following reason. It requires reactivation of a verbal element, just like with Gapping, but it is the whole VP (e.g., *ordered coffee and tea at the cafe* in Table 7.2 (c); *drank coffee and tea at the cafe* in Table 7.2 (d)) that is retrieved at the ellipsis region. Consequently, VPE has no remnant direct object NP in the ellipsis clause, and thus the two VPE baseline conditions are not expected to induce any effect of plausibility. Therefore, having VPE as a baseline allows us to safely assert that any reading-time differences between the Gapping conditions at the regions of interest are due to the successful

reconstruction of a verb gap in the gapped clause and not due to the difference in verbs in the non-gapped clause.

Following the two studies by Kaan and her colleagues (2004, 2013), this study employed proper names as subjects. This was done to make the two clausal conjuncts in a sentence as parallel as possible and to encourage an interpretation of *and* between the two clauses as coordinating two clauses rather than conjoining an object NP in the first clause and a subject NP in the second clause. On the other hand, the objects in both conjuncts were always a full NP referring to an inanimate object. Care was taken to ensure that the verbs in the Plausible and Implausible conditions were comparable in terms of frequency in the Corpus of Contemporary American English (COCA; Davies, 2008–). The result of a *t*-test did not reveal a significant difference in frequency between the Plausible and Implausible verb sets[2] ($t(38) = -0.879$, $p = .385$, Cohen's $d = 0.278$).

The filler sentences, which were designed to be similar to the critical sentences in terms of length and complexity, involved (a) Right Node Raising ($k = 20$; e.g., *Bill made [e] and John sold pizza and pasta during the vacation, according to their mother*); (b) subject–verb number agreement ($k = 20$; e.g., *The fires in the apartment were caused by a cigarette butt thrown on the carpet*); (c) *where*-clause ($k = 6$; e.g., *Sara drew the spider and the ant in the park where the car and the truck were parked*); and (d) *Do So* Anaphora ($k = 4$; e.g., *Robin read a book on the couch and Leslie did so on the bench*). The subject–verb number agreement type fillers were slightly modified versions of the sentences used by Jiang (2004). Approximately 56% of the fillers were plausible/grammatical, while the remaining were implausible/ungrammatical.

All the experimental and filler sentences were followed by yes/no comprehension questions to ensure that participants would pay attention to the sentences and parse them for meaning. The questions for the critical sentences always targeted the first clause to minimize the likelihood of participants becoming consciously aware of the verb gap in the second clause. The comprehension question following the item in Table 7.2 is shown in (1).

[2] The plausible and implausible verb sets were: (1) *order–drink*, (2) *have–spill*, (3) *find–break*, (4) *buy–design*, (5) *heat–bake*, (6) *serve–cook*, (7) *clean–close*, (8) *replace–open*, (9) *paint–empty*, (10) *grab–swallow*, (11) *discover–wash*, (12) *observe–eat*, (13) *prepare–peel*, (14) *draw–kill*, (15) *drop–cut*, (16) *receive–rip*, (17) *own–drive*, (18) *admire–water*, (19) *sell–play*, (20) *enjoy–watch*.

(1) Question: Did Bill order milk?

Answer options:

1. yes
2. no

7.2.3 Norming study.

A norming study was conducted to confirm, first, that the Gapping-Plausible and Gapping-Implausible condition sentences were judged as expected and, second, that the VPE-Plausible and VPE-Implausible condition sentences were roughly equal in terms of plausibility.

7.2.3.1 Participants.

A separate group of 16 English native speakers, who were undergraduate students at the University of Hawai‘i at Mānoa, participated in the norming study online via Ibex Farm (Drummond, 2007).

7.2.3.2 Materials and procedure.

The norming task had the 20 critical sentences and the 50 filler sentences that were used in the aforementioned self-paced reading task. The critical sentences were distributed in four lists. Participants were asked to rate these sentences on a scale ranging from 1 (definitely not plausible) to 4 (definitely plausible). There was also a fifth option labeled as ‘I don’t know’. This norming task was designed and presented on Ibex Farm (Drummond, 2007). The order of sentences was pseudo-randomized for each participant and no two sentences from the same condition were presented consecutively. The sentences were displayed in black letters with Times New Roman font (12pt) and a white background.

7.2.3.3 Results.

After removing ‘I don’t know’ responses (5.00% of the data), ratings for all sentences including fillers were transformed to *z*-scores. Using a *z*-score transformation reduces the impact of extreme data points and varying uses of the interval scale by participants (Casasanto, Hofmeister, Sag, Ohlsson, & Catrambone, 2010) and normalizes the data.

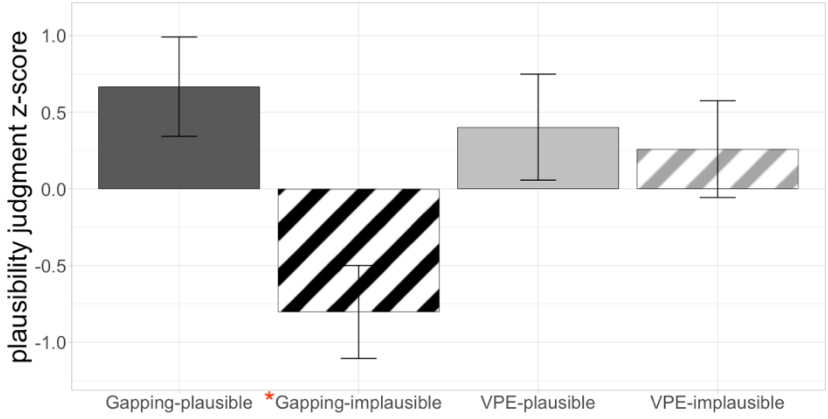
To ascertain any effects of *Construction* (Gapping; VPE) and *Plausibility* (Plausible; Implausible) in the plausibility judgments of the participants, a linear mixed-effects regression model was constructed with *Construction* and *Plausibility* as fixed effects (contrast-coded with [−.5, .5]) and with *participant* and *item* as random effects. As shown in Table 7.3 and Figure 7.1, the model output revealed a main effect of *Plausibility* ($\beta = 0.822$, $SE = 0.235$, $p = .002$) and a main interaction between *Construction* and *Plausibility* ($\beta = -1.343$, $SE = 0.383$, $p = .002$). There was no significant effect of *Construction* ($\beta = 0.408$, $SE = 0.274$, $p = .155$).

Table 7.3
Results of the Linear Mixed-Effects Regression for the Plausibility Judgment Data in Study 4

	Estimate (β)	Standard Error (SE)	p
(Intercept)	0.106	0.173	.547
Construction	0.408	0.274	.155
Plausibility	0.822	0.235	.002
Construction × Plausibility	−1.343	0.383	.002

Note. Model formula: lmer(Judgment ~ Construction * Plausibility + (1 + Construction * Plausibility | participant) + (1 + Construction * Plausibility | item))

Figure 7.1
Mean Plausibility Judgment Ratings per Condition in Study 4



Note. The error bars indicate 95% confidence intervals.

Importantly, the follow-up pairwise comparison showed significantly higher ratings in the Gapping-Plausible condition than in the Gapping-Implausible condition ($\beta = 1.493$, $SE = 0.382$, $p < .001$). In contrast, there was no significant difference in plausibility ratings between the two VPE conditions ($\beta = 0.133$, $SE = 0.173$, $p = .449$). The sentences of the Gapping-Plausible condition ($M = 0.67$; $SD = 1.41$), the VPE-Plausible condition ($M = 0.40$; $SD = 1.47$), and the

VPE-Implausible condition (M = 0.26; SD = 1.39) obtained plausibility rating z-scores higher than zero, which was the middle point of the scale, whereas the sentences in the Gapping-Implausible condition earned z-scores below zero (M = −0.80; SD = 1.33). Overall, these findings indicate that only the Gapping-Implausible condition was perceived as implausible by English native speakers, as intended.

7.2.4 Procedure.

All participants first completed a language background questionnaire asking for their native language, age of English onset,[3] and other language-related information (see Appendix E). Next, they undertook a noncumulative moving-window self-paced reading task (Just, Carpenter, & Woolley, 1982). The sentences were presented in a segment-by-segment fashion in black letters (Times New Roman, 12pt) on a white background in the center of the screen. All items appeared on a single line. Participants were first given written instructions which told them to read each segment as quickly as possible for comprehension and to press the spacebar to move to the next segment. Next, they completed a practice session with five items and then moved on to the experimental session, which consisted of 20 critical sentences plus 50 fillers presented in pseudo-randomized order. Each item was followed by a yes/no comprehension question that appeared in the center of the screen with the answer choices beneath it. For half of the questions, the correct answer was "1. yes" and for the other half, it was "2. no." Finally, all participants completed a cloze test as a measure of English proficiency. All of these tasks were presented on Ibex Farm (Drummond, 2007). The total duration of the experimental session was approximately 50 minutes.

7.3 Predictions

Based on Kaan et al.'s (2004) results, the English native speakers were expected to show processing difficulty at Segment 8 and/or Segment 9 in the Gapping-Implausible condition (vs. the Gapping-Plausible condition) where the gapped verb and its following object NP make an implausible combination. In contrast, they were not expected to show any plausibility effect for the VPE baseline conditions. This should thus result in an interaction effect between *Construction* and *Plausibility* at Segment 8 and/or Segment 9. It was also expected that

[3] Participants were asked to write "0" on this item if they were English native speakers.

follow-up pairwise comparisons at these regions would reveal a significant difference between the reading times (RTs) in the Gapping conditions—resulting from longer RTs in the Gapping-Implausible condition—but no significant difference between the RTs in the VPE conditions.

We would expect the L2ers as a group to display the same processing patterns as the English native speakers if and only if they possess implicit knowledge of Gapping and are able to use syntactic information related to Gapping during real-time processing. If however the L2ers do not show the target-like pattern in the by-group analysis, then it may still be possible to observe proficiency effects such that more advanced L2ers show the target-like RT pattern and less advanced ones do not. This outcome would indicate that L2ers with higher proficiency are able to identify verb gaps and retrieve the verb information in Gapping sentences.

7.4 Data Analysis

Prior to analyzing the RT data, participants with lower than 80% accuracy on the comprehension questions were removed from the data set. As previously mentioned, this led to 12 English native speakers and 11 L2ers being excluded.

Next, individual trials with incorrect answers to the comprehension questions were excluded from further analyses (L1-English: 10.85%; L2-English: 14.79%). Following Keating and Jegerski (2015) and Keating, Jegerski, & VanPatten (2016), the raw RTs were screened for extreme values and all RTs below 100ms or above 4,000ms were removed. This affected 0.16% of the L1-English data set and 0.44% of the L2-English data set. Outlier values, which were defined as RTs in excess of 1.5 standard deviations above or below the mean for a given condition, were identified both by participant (L1-English: 4.58%; L2-English: 5.05%) and by item (L1-English: 7.63%; L2-English: 6.56%), and were removed. Then residual RTs (henceforth, RRTs) were computed in order to adjust for differences in word length and participants' reading rates (Ferreira & Clifton, 1986). For this procedure, a linear mixed-effects model was first fitted to the raw RTs with *number of characters* as a fixed effect and *participants* as a random effect.[4] Then the RTs predicted by each participant's regression equation were subtracted from the raw RTs to obtain RRTs for the data analyses (e.g., Roberts & Liszka, 2020).

[4] Model formula: lmer(RT ~ number of characters + (1 + number of characters | participant))

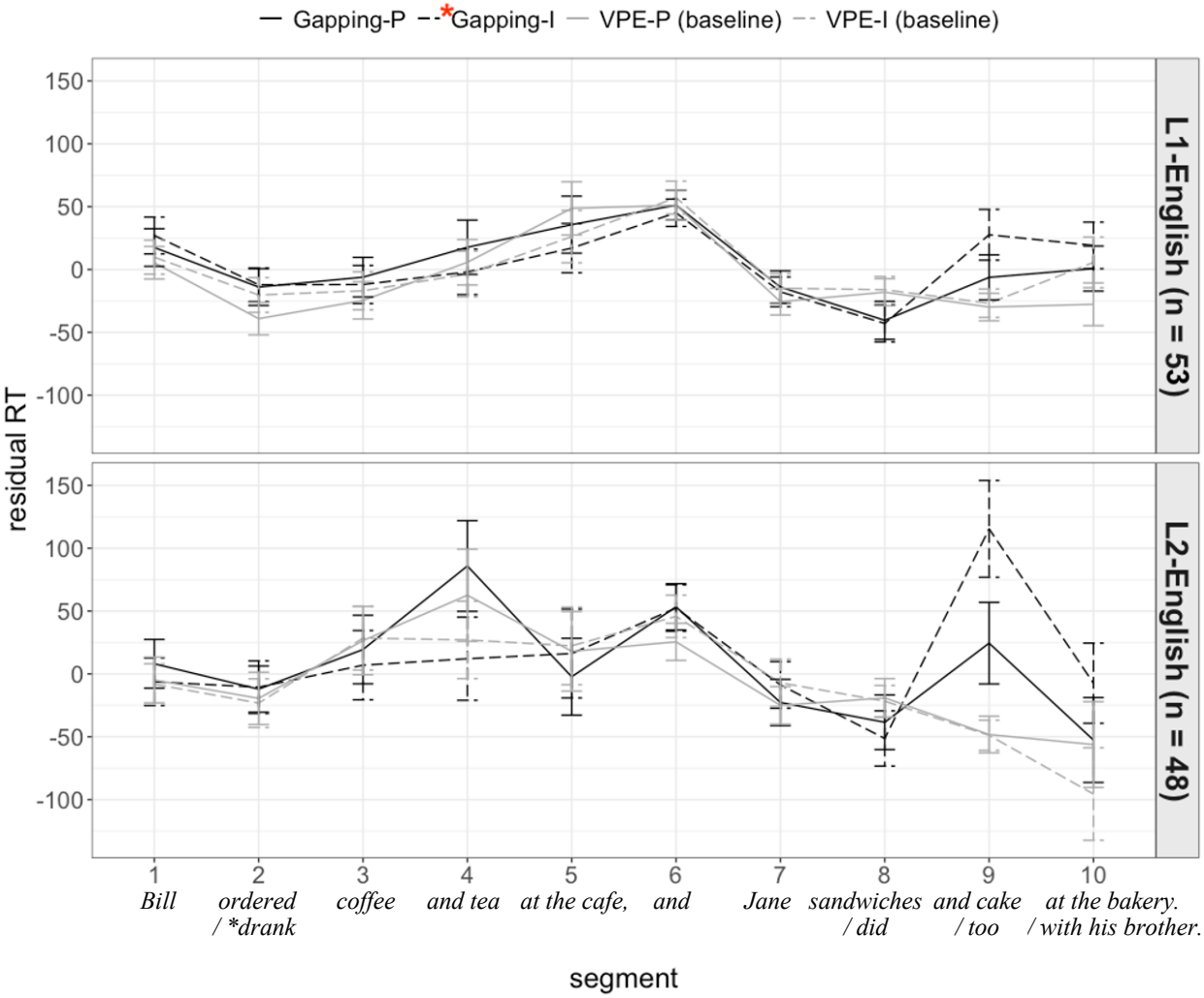
To test for any effects of *Construction* (Gapping; VPE) and *Plausibility* (Plausible; Implausible) in the RRTs for each group, a linear mixed-effects regression model was fitted to the RRTs in each of the critical segments with *Construction* and *Plausibility* as fixed effects (contrast-coded with [−.5, .5]) and with *participant* and *item* as random effects. The L2 data were further examined for a potential effect of *Proficiency* by constructing a separate mixed-effects model in which *Proficiency* was included as a continuous fixed effect in addition to the fixed effects *Construction* and *Plausibility* (contrast-coded with [−.5, .5]) and the random effects *participant* and *item*. In the case that an interaction was found, additional mixed-effects regression analyses were then conducted for pairwise comparisons. For all of the mixed-effects regression models, I constructed the maximal random effects structure allowed by the design (Barr et al., 2013; Matuschek et al., 2017; Stroup, 2012) in R (R Core Team, 2018). The model formula for each analysis is reported in the corresponding results table.

7.5 Results

The participants' overall accuracy on the comprehension questions was 89.33% (*SD* = 4.30) for the 53 native speakers and 87.25% (*SD* = 4.44) for the 48 L2ers. These high scores show that both groups were paying attention to the self-paced reading task.

Figure 7.2 displays the mean RRTs for each group per segment and per condition. (The mean raw RTs for each group per segment and per condition are reported in Appendix M.) At Segment 9, both the English native speakers and the L1-Korean L2ers showed longer RRTs for the Gapping-Implausible condition than for the Gapping-Plausible condition, but such plausibility effects were absent from the VPE baseline conditions, as predicted. Neither group showed any noticeable processing patterns at Segment 8.

Figure 7.2
Mean Residual Reading Times (in ms) per Segment, Condition, and Group



Notes. The top panel shows the results for the English native speakers (L1-English), and the bottom panel shows the results for the L1-Korean L2ers of English (L2-English). Gapping-P = Gapping-Plausible, *Gapping-I = Gapping-Implausible; VPE-P = VPE-Plausible, VPE-I = VPE-Implausible. The error bars indicate 95% confidence intervals.

In the next section, I first report the results of the linear mixed-effects regression analysis on the RRTs at Segment 8 (the critical region). Then I report the results obtained from the analysis of the RRTs at Segment 9 (the spill-over region).

7.5.1 Reading times at the critical region (Segment 8).

The mixed-effects model fitted to the L1-English data at Segment 8 revealed a significant effect of *Construction* ($\beta = 25.575$, $SE = 9.073$, $p = .010$), as shown in Table 7.4. This effect

resulted from longer RTs in the Gapping conditions relative to the VPE conditions, thus indicating that the English native speakers identified a verb gap. However, there was neither a significant effect of *Plausibility* ($\beta = 0.352$, $SE = 6.475$, $p = .957$) nor a significant interaction between *Construction* and *Plausibility* ($\beta = -4.141$, $SE = 15.818$, $p = .796$).

Table 7.4
Results of the Linear Mixed-Effects Regression at the Critical Region (Segment 8) for the English Native Speaker Data in Study 4

	Estimate (β)	Standard Error (SE)	p
(Intercept)	−29.060	5.037	< .001
Construction	25.575	9.073	.010
Plausibility	0.352	6.475	.957
Construction × Plausibility	−4.141	15.818	.796

Note. Model formula: lmer(Residual RT ~ Construction * Plausibility + (1 + Construction * Plausibility | participant) + (1 + Construction * Plausibility | item))

The model for the L2 data also showed a significant effect of *Construction* at Segment 8 ($\beta = 22.596$, $SE = 10.251$, $p = .032$) with longer RTs in the Gapping conditions than the VPE conditions (see Table 7.5). This indicates that the L2ers detected the presence of the verb gap.

Table 7.5
Results of the Linear Mixed-Effects Regression at the Critical Region (Segment 8) for the L2 Data in Study 4

	Estimate (β)	Standard Error (SE)	p
(Intercept)	−30.768	5.331	< .001
Construction	22.596	10.251	.032
Plausibility	4.577	14.518	.755
Construction × Plausibility	−4.941	24.061	.839

Note. Model formula: lmer(Residual RT ~ Construction * Plausibility + (1 + Construction * Plausibility | participant) + (1 + Construction * Plausibility | item))

However, the L2 data at Segment 8 had no significant effect of *Plausibility* ($\beta = 4.577$, $SE = 14.518$, $p = .755$) or significant interaction between *Construction* and *Plausibility* ($\beta = -4.941$, $SE = 24.061$, $p = .839$).

In order to investigate the effect of *Proficiency*, a separate linear mixed-effects model was constructed with the continuous factor *Proficiency* added. As shown in Table 7.6, there was no significant effect of *Proficiency* ($\beta = 0.147$, $SE = 0.551$, $p = .790$). Nor was there a significant

interaction between *Construction* and *Proficiency* (β = 0.232, *SE* = 1.070, p = .829) or between *Plausibility* and *Proficiency* (β = 0.204, *SE* = 1.107, p = .854). A three-way interaction of *Construction*, *Plausibility*, and *Proficiency* did not emerge, either (β = 1.376, *SE* = 1.991, p = .492).

Table 7.6
Results of the Linear Mixed-Effects Regression at the Critical Region (Segment 8) for the L2 Data in Study 4 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	−35.282	17.646	.051
Construction	15.349	34.271	.656
Plausibility	−1.801	36.814	.961
Proficiency	0.147	0.551	.790
Construction × Plausibility	−47.224	65.549	.473
Construction × Proficiency	0.232	1.070	.829
Plausibility × Proficiency	0.204	1.107	.854
Construction × Plausibility × Proficiency	1.376	1.991	.492

Note. Model formula: lmer(Residual RT ~ Construction * Plausibility * Proficiency + (1 + Construction * Plausibility | participant) + (1 + Construction * Plausibility | item))

7.5.2 Reading times at the spill-over region (Segment 9).

As shown in Table 7.7, the mixed-effects model for the L1-English data showed a significant effect of *Construction* (β = −39.033, *SE* = 7.501, p < .001), with longer RRTs in the Gapping conditions vs. the VPE conditions. There was also a significant effect of *Plausibility* (β = −18.067, *SE* = 7.726, p = .021). Importantly, the model revealed a significant interaction of *Construction* and *Plausibility* (β = 30.644, *SE* = 15.009, p = .042). This interaction resulted from the longer RRTs in the Gapping-Implausible condition than in the Gapping-Plausible condition (β = −36.717, *SE* = 14.892, p = .017). There was no significant difference between the two VPE conditions (β = −3.056, *SE* = 7.738, p = .693). These results indicate that the English native speakers experienced processing difficulty at the spill-over region only when the gapped verb was implausible for the following direct object. This suggests that they identified and resolved the verb gap.

Table 7.7
Results of the Linear Mixed-Effects Regression at the Spill-Over Region (Segment 9) for the English Native Speaker Data in Study 4

	Estimate (β)	Standard Error (SE)	p
(Intercept)	−10.370	5.525	.073
Construction	−39.033	7.501	< .001
Plausibility	−18.067	7.726	.021
Construction × Plausibility	30.644	15.009	.042

Note. Model formula: lmer(Residual RT ~ Construction * Plausibility + (1 + Plausibility | participant) + (1 | item))

The model for the L2ers' data showed a significant effect of *Construction* (β = −122.652, *SE* = 23.405, p < .001) and *Plausibility* (β = −38.915, *SE* = 14.437, p = .012), as shown in Table 7.8.

Table 7.8
Results of the Linear Mixed-Effects Regression at the Spill-Over Region (Segment 9) for the L2 Data in Study 4

	Estimate (β)	Standard Error (SE)	p
(Intercept)	3.495	11.235	.758
Construction	−122.652	23.405	< .001
Plausibility	−38.915	14.437	.012
Construction × Plausibility	92.836	39.857	.028

Note. Model formula: lmer(Residual RT ~ Construction * Plausibility + (1 + Construction * Plausibility | participant) + (1 + Construction * Plausibility | item))

Crucially, there was a significant interaction between *Construction* and *Plausibility* (β = 92.836, *SE* = 39.857, p = .028). The follow-up pairwise comparisons revealed that this interaction stemmed from the fact that the L2ers had significantly longer RRTs for the Gapping-Implausible condition than the Gapping-Plausible condition (β = −88.873, *SE* = 29.864, p = .008). However, no significant difference was found between the VPE-Plausible condition and the VPE-Implausible condition (β = 0.209, *SE* = 12.478, p = .987).

A *Proficiency* effect was investigated using a separate mixed-effects model with the continuous fixed effect *Proficiency* added. As shown in Table 7.9, this model did not reveal a significant effect of *Proficiency* (β = −0.274, *SE* = 0.724, p = .707).

Table 7.9
Results of the Linear Mixed-Effects Regression at the Spill-Over Region (Segment 9) for the L2 Data in Study 4 with the Factor Proficiency *Added*

	Estimate (β)	Standard Error (SE)	p
(Intercept)	11.739	24.723	.637
Construction	−194.629	51.437	< .001
Plausibility	−25.545	41.200	.538
Proficiency	−0.274	0.724	.707
Construction × Plausibility	−22.828	97.286	.816
Construction × Proficiency	2.350	1.509	.128
Plausibility × Proficiency	−0.447	1.262	.724
Construction × Plausibility × Proficiency	3.777	2.916	.203

Note. Model formula: lmer(Residual RT ~ Construction * Plausibility * Proficiency + (1 + Construction * Plausibility | participant) + (1 + Construction * Plausibility | item))

The factor *Proficiency* did not significantly interact with *Construction* (β = 2.350, *SE* = 1.509, p = .128) or *Plausibility* (β = −0.447, *SE* = 1.262, p = .724). The three-way interaction of *Construction*, *Plausibility*, and *Proficiency* was not significant (β = 3.777, *SE* = 2.916, p = .203).

In sum, the RRTs of the English native speakers and the L2ers were significantly longer for the Implausible condition than the Plausible condition in the Gapping construction (but not in the VPE construction). In the L2ers' processing patterns, no proficiency effects were found. These results suggest that the L2ers, like the English native speakers, posited a verb gap and integrated the NP object with the gapped verb using the verb information available at the gap site.

7.6 Discussion

The aim of this study was to investigate whether L2ers are able to identify a verb gap and reconstruct verb information at the gap site in the processing of Gapping sentences. The L1-Korean L2ers of English in this study showed longer RRTs in the Gapping-Implausible condition than in the Gapping-Plausible condition at the spill-over region, just as the native speaker controls did. I interpret this finding as indicating that the L2ers identify and process verb gaps in a target-like manner.

Successful verb-gap processing has been found in native processing studies on English Gapping (see §3.4). For example, N. Kim et al.'s (2020) eye-tracking-while-reading study found that English native speakers postulated a verb gap at the region that was potentially compatible

with a Gapping structure; while reading the sentence *The guitarist hid behind the curtain suddenly, and the singer behind the stage hid from the sneaky photographers*, English native speakers showed a slowdown at the region *hid from the sneaky photographers* presumably because they constructed a Gapping structure for the NP-PP sequence in the second clause (i.e., *the singer [e] behind the stage*) and needed to reanalyze it as subject of the second clause. Kaan et al.'s (2004) ERP study showed that English native speakers detected a verb gap at the gap site and attempted to integrate the gapped verb with the following object. Using a self-paced reading paradigm, the current study replicated the previous finding of English native speakers' successful verb-gap processing and furthermore extended it to L2 processing.

It should be noted that this study did not independently test L2ers' knowledge of Gapping. However, I showed in Study 2 (see Chapter 5) and Study 3 (see Chapter 6) that a separate group of advanced L1-Korean L2ers of English do have target knowledge of Gapping, in terms of both grammaticality and interpretation. These results from the acquisition studies, together with the current processing results, constitute converging evidence that (ultimately) L2ers, like native speakers, are able to represent and process verb gaps.[5]

Interestingly, it was at the spill-over region, and not the critical region, that the English native speakers and the L2ers displayed the predicted processing patterns. This can be explained in two ways. First, it may indicate delayed processing, which is attributable to the complex structure of Gapping. As parsers had to identify the verb gap and reconstruct the meaning of the Gapped conjunct by retrieving the verbal information, processing Gapping may have been hard for the parsers. Alternatively, the spill-over effects may have been induced by a "task-induced button-press rhythm" (e.g., Koornneef & Van Berkum, 2006). The participants might have pressed the space bar quickly to see the next region, and this rapid performance in turn might explain why the processing effect was observed in the segment following the critical region rather than on the critical region itself. It is thus conceivable that the absence of the predicted RRTs at the critical region did not result from delayed processing on the part of the English native speakers and the L2ers.

Note, however, that unexpected patterns were also found at other regions. For instance, at Segment 2, a significant effect of *Construction* ($\beta = -15.987$, $SE = 7.138$, $p = .030$) was found in

[5] I am grateful to Amy Schafer for a helpful discussion about this point (personal communication, 6 March 2020).

the mixed-effects model for the L1-English data. Given that this region had the same set of verbs for the Gapping-Plausible and VPE-Plausible conditions (e.g., *ordered*) and the same set of verbs for the Gapping-Implausible and VPE-Implausible conditions (e.g., *drank*), it is hard to explain the underlying source of this effect. In addition, the L1-English model at Segment 10 showed a main effect of *Plausibility* ($\beta = -24.846$, $SE = 11.679$, $p = .047$), with longer RRTs for the Implausible conditions than the Plausible conditions. The reason for this result is also unclear because this region had the same set of PPs for the Gapping conditions (e.g., *at the bakery*) and the same set of PPs for the VPE conditions (e.g., *with his brother*) regardless of *Plausibility*. The model for the L2 data also revealed an unexpected result: There was a significant effect of *Proficiency* at Segment 3 ($\beta = 35.942$, $SE = 2.751$, $p = .009$) such that higher-proficiency L2ers showed shorter RRTs. It is hard to pinpoint what led to this processing pattern. (No other significant effects or interactions were found from the analyses at Segments 1–10.)

Let's return to our findings at Segment 9 (the spill-over region). One might suggest that the plausibility effects stemmed from some sort of non-syntactic combination of lexical items (for related discussion, see also Kaan et al., 2013). One of the filler types, the *Where*-Plausible vs. *Where*-Implausible controls ($k = 6$), was included specifically to address this possibility. The plausible-control and implausible-control versions of these sentences differed only with regard to the overt verb in the first clause, as illustrated in (2a) and (2b). The regions of interest were the same as those in the critical conditions. The critical region was the first NP conjunct within the *where*-clause (e.g., *sandwiches* in (2)), and the following conjunction and NP (e.g., *and cake* in (2)) served as the spill-over region.

(2) a. Where-Plausible (control)
Henry / ordered / coffee / and tea / at the cafe / where / sandwiches / and cake / were / very popular.

b. Where-Implausible (control)
Henry / drank / coffee / and tea / at the cafe / where / sandwiches / and cake / were / very popular.

A linear mixed-effects regression model was constructed for each group to ascertain whether there were any effects of *Plausibility* in the RRTs. This model included *Plausibility*

(contrast-coded with [−.5, .5]) as fixed effects and *participant* and *item* as random effects. To ascertain any effects of *Proficiency* in the L2 data, I built an additional model including *Plausibility* and *Proficiency* as fixed effects and *participant* and *item* as random effects. Both the L1-English model and the L2-English model showed no significant effect of *Plausibility* at either the critical region (L1-English: $\beta = 3.676$, $SE = 11.971$, $p = .762$; L2-English: $\beta = -21.323$, $SE = 24.297$, $p = .408$) or the spill-over region (L1-English: $\beta = -11.992$, $SE = 13.540$, $p = .405$; L2-English: $\beta = 16.243$, $SE = -35.202$, $p = .658$). In addition, the L2-English model including the factor *Proficiency* did not reveal a significant interaction between *Plausibility* and *Proficiency* at either of the regions of interest (critical: $\beta = 1.402$, $SE = 1.775$, $p = .433$; spill-over: $\beta = 1.250$, $SE = 2.098$, $p = .554$). These results suggest that non-syntactic combination of lexical items cannot explain the plausibility effects that were found between the two Gapping conditions in the current study.[6]

7.7 Conclusion

All in all, this study found that L1-Korean adult L2ers with, on average, relatively high English proficiency are able to identify verb gaps and quickly reconstruct verbal information after arriving at the gap site. This finding provides further evidence for an account of L2 development in which L2ers can come to make use of syntactic information during online processing in the same way native language users do.

[6] Kamil Deen (personal communication, 17 April 2020) pointed out that the L2ers' results on the *where*-fillers in (2) do not fully eliminate the possibility that their target-like processing of Gapping came from non-syntactic lexical considerations. Specifically, when reading Gapping sentences, the L2ers might have experienced processing breakdown due to encountering two consecutive NPs in the gapped clause, i.e., the subject NP followed by the object NP (e.g., *Jane [e] sandwiches and cake* in *Bill ordered/*drank coffee and tea at the cafe, and Jane [e] sandwiches and cake at the bakery.*), at which point it might have been possible for them to associate the object NP with the overt verb without constructing a verb gap. To address this issue, Bonnie D. Schwartz (personal communication, 17 April 2020) suggested that I recruit more L1-Korean L2ers with diverse proficiency levels to take part in the self-paced reading task and the offline task testing grammatical knowledge of Gapping. If only the higher-proficiency L2ers (but not the lower-proficiency L2ers) display both target knowledge of Gapping and target-like processing patterns, this would buttress the claim that higher-proficiency L2ers represent and process verb gaps syntactically.

CHAPTER VIII
GENERAL DISCUSSION AND CONCLUSION

This chapter begins by summarizing the main findings of this dissertation (§8.1). It then provides the combined results of the two acquisition studies (§8.2). Next, it discusses the theoretical implications of the findings in this dissertation, focusing on the learnability problems associated with the grammaticality and interpretation contrasts between VP-Ellipsis (VPE) and Gapping (§8.3.1), the mechanisms underlying adult L2 acquisition (§8.3.2), and the parsing strategies that shape L2 sentence processing (§8.3.3). Section 8.4 concludes the dissertation.

8.1 Summary of Main Findings

This dissertation investigated the grammaticality and interpretation contrasts between VPE and Gapping in L1 and L2 acquisition. Crucially, these contrasts involve learnability problems for both L1-English children and L1-Korean L2ers of English (see §2.4). Study 1 used a natural language processing analysis to demonstrate that the input these learners receive is insufficient for acquiring the target contrasts. Studies 2 and 3 tested whether the participants possessed implicit knowledge of the contrasts by means of an acceptability judgment task (AJT) and a sentence-picture matching task, with results showing that the L1-English children evinced the grammaticality contrast as early as age 5;11 and the interpretation contrast as early as age 5;6. The early L2ers with higher proficiency and most of the late L2ers had also developed both contrasts. Finally, Study 4 demonstrated via a self-paced reading task that adult L2ers and L1 adult controls were both able to successfully identify and resolve the verb gap while processing Gapping sentences in real time.

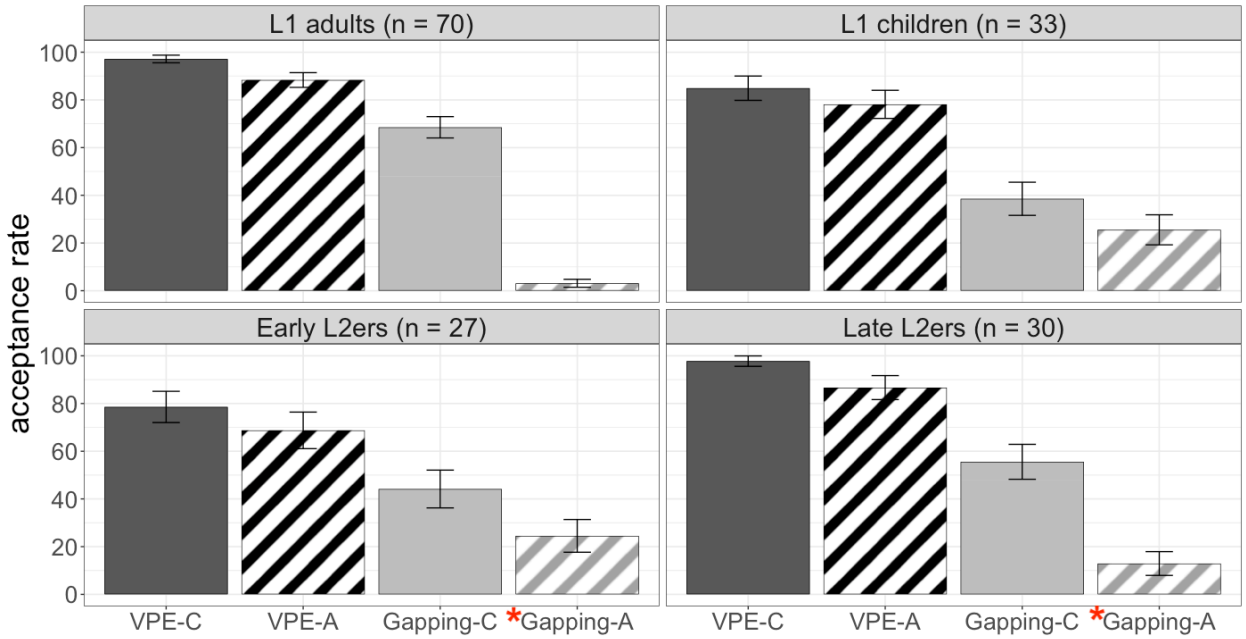
8.2 Combined Results of the Two Acquisition Studies.

This section combines and discusses the results from the acceptability judgment study (Study 2) and the interpretation study (Study 3). I begin this section by summarizing the group results for the L1 adults, L1 children, early L2ers, and late L2ers. The L1 adults showed clear differences between VPE and Gapping for both the grammaticality contrasts (see Figure 8.1) and the interpretation contrasts (see Figure 8.2) contrasts. They accepted VPE in a conjunct clause, VPE in an adjunct clause, and Gapping in a conjunct clause but rejected (ungrammatical)

Gapping in an adjunct clause. They also allowed VPE with a subject reading and Gapping with an object reading but rejected VPE with an object reading. Although they allowed Gapping with a subject reading only about half the time, this pattern, I argued, can be reasonably viewed as a consequence of processing difficulty (see §6.4.1).

Figure 8.1

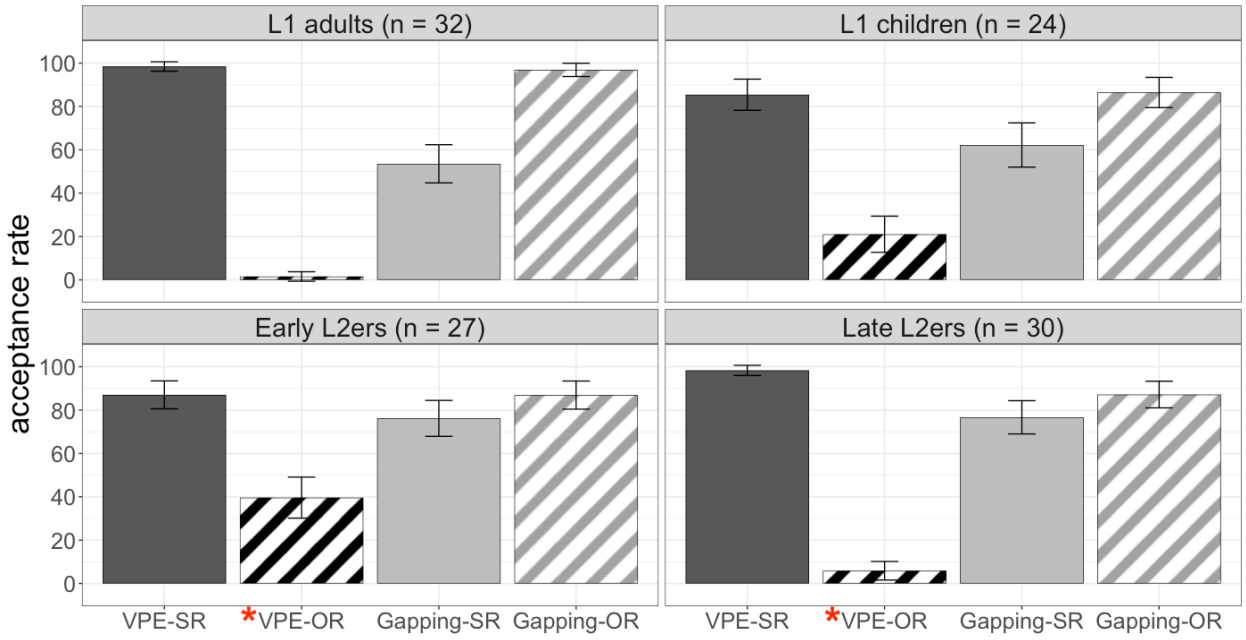
Mean Acceptance Rates (in %) in the Acceptability Judgment Task per Condition and Group



Notes. Reproduced from Figure 5.2. VPE-C: VPE in a conjunct clause; VPE-A: VPE in an adjunct clause; Gapping-C: Gapping in a conjunct clause; *Gapping-A: Gapping in an adjunct clause. Error bars represent 95% confidence intervals.

Figure 8.2

Mean Acceptance Rates (in %) in the Picture-Sentence Matching Task per Condition and Group



Notes. Reproduced from Figure 6.2. VPR-SR: VPE with a subject reading; *VPE-OR: VPE with an object reading; Gapping-SR: Gapping with a subject reading; Gapping-OR: Gapping with an object reading. Error bars represent 95% confidence intervals.

The late L2ers, many of whom had high proficiency scores, displayed the same pattern of target grammaticality and interpretation contrasts that the L1 adults did. Although the L1 children and early L2ers (as groups) did not fully reject Gapping in an adjunct clause (see Figure 8.1) or VPE with an object reading (see Figure 8.2), their acceptance rates for these illicit conditions tended to decrease as their age increased in the case of the L1 children (see §5.3.2, §6.3.2) and as their proficiency increased in the case of the early L2ers (see §5.3.3, §6.3.3).

In order to try to understand the developmental sequences involved in the grammaticality and interpretation contrasts between VPE and Gapping, I further inspected the judgments of individual learners using *Age* as a guideline for L1 children and *Proficiency* as a guideline for L2ers. Individual participants were marked for the presence of knowledge involved in the contrasts at issue based on the strict criteria that I created for Study 2 and Study 3. Recall that the four criteria established for the grammaticality contrast were (a) knowledge of VPE, (b) knowledge of the ungrammaticality of backward Gapping, (c) knowledge of the ungrammaticality of Gapping in adjunct clauses, and (d) knowledge of the contrast between Gapping in conjunct clauses and Gapping in adjunct clauses (see §5.4.2). The four criteria set for

the interpretation contrast were (a) knowledge of the possibility of VPE with a subject reading, (b) knowledge of the possibility of Gapping with an object reading, (c) knowledge of the contrast between VPE with a subject reading and VPE with an object reading, and (d) knowledge of the impossibility of VPE with an object reading (see §6.4.2). Participants were given a check mark (✓) for each criterion they passed. I used shading to mark participants who passed all four criteria for the contrast at issue. Importantly, developmental sequences were identified based on the following requirements: (a) Each developmental stage should start with at least two consecutive learners (sorted by *Age* for the L1 children and by *Proficiency* for the L2ers) who passed the criterion under inspection and (b) no stage should have more than two consecutive learners who fail to pass the criterion. The developmental sequences identified for the L1 children, the early L2ers, and the late L2ers are shown in, respectively, Tables 8.1–8.3.

Table 8.1

*Developmental Sequence of the Grammaticality and Interpretation Contrasts between VPE and Gapping: L1 Children in Chronological Order (*n *= 24–33)*

		Grammaticality contrast				Interpretation contrast				
Age	**PPT**	**Knowledge of the grammaticality of VPE**	**Knowledge of the ungrammaticality of *backward Gapping**	**Knowledge of the ungrammaticality of *Gapping-A**	**Knowledge of the contrast between Gapping-C and *Gapping-A**	**Knowledge of the possibility of VPE-SR**	**Knowledge of the possibility of Gapping-OR**	**Knowledge of the contrast between VPE-SR and *VPE-OR**	**Knowledge of the impossibility of *VPE-OR**	**Developmental stage**
3;3	L1C_04		✓			N/A	N/A	N/A	N/A	
3;11	L1C_05	✓			✓	N/A	N/A	N/A	N/A	
4;0	L1C_06			✓	✓	N/A	N/A	N/A	N/A	
5;0	L1C_03					N/A	N/A	N/A	N/A	
5;4	L1C_10	✓			✓	✓	✓			(1) Knowledge of the grammaticality of VPE and the possibility of VPE-SR and Gapping-OR
5;6	L1C_28	✓	✓			✓	✓	✓	✓	(2) Knowledge of the contrast between VPE-SR and *VPE-OR
5;8	L1C_25	✓				✓	✓	✓		
5;10	L1C_13	✓	✓	✓		✓		✓	✓	(3) Knowledge of the ungrammaticality of *backward Gapping and the impossibility of *VPE-OR
5;11	L1C_14		✓		✓		✓	✓	✓	
5;11	L1C_26	✓	✓	✓	✓	✓	✓	✓		
6;1	L1C_20	✓						✓	✓	
6;3	L1C_17	✓		✓		✓	✓	✓		(4) Knowledge of the ungrammaticality of *Gapping-A
6;3	L1C_24	✓	✓	✓		✓	✓	✓	✓	
6;6	L1C_31	✓				✓	✓	✓		
6;6	L1C_15	✓	✓	✓	✓	✓	✓	✓	✓	(5) Knowledge of the contrast between Gapping-C and *Gapping-A
6;7	L1C_18	✓		✓	✓	N/A	N/A	N/A	N/A	
6;8	L1C_07	✓			✓	✓	✓			
6;8	L1C_12	✓	✓	✓		✓	✓	✓	✓	
6;8	L1C_27	✓			✓	✓	✓	✓	✓	
6;9	L1C_33	✓			✓	✓	✓	✓	✓	
6;9	L1C_23	✓	✓	✓		✓	✓	✓	✓	
6;10	L1C_16	✓			✓	✓	✓	✓	✓	
6;10	L1C_30	✓	✓		✓	✓	✓	✓	✓	
6;11	L1C_09	✓	✓	✓	✓		✓			
6;11	L1C_11		✓	✓	✓	✓	✓	✓	✓	
6;11	L1C_21	✓	✓	✓	✓	N/A	N/A	N/A	N/A	
7;1	L1C_22	✓	✓	✓		✓	✓	✓	✓	
7;1	L1C_32	✓	✓	✓		✓	✓	✓	✓	
7;2	L1C_29	✓		✓	✓	✓	✓	✓	✓	
7;2	L1C_19	✓	✓	✓			✓	✓	✓	
7;7	L1C_01	✓	✓	✓	✓	N/A	N/A	N/A	N/A	
7;8	L1C_08	✓	✓			N/A	N/A	N/A	N/A	
7;9	L1C_02	✓	✓	✓	✓	N/A	N/A	N/A	N/A	

Notes. The check mark (✓) indicates presence of knowledge. The four criteria for the grammaticality contrast were as follows: knowledge of the grammaticality of VPE (8–12 correct out of 12); knowledge of the ungrammaticality of *backward Gapping (3 correct out of 3); knowledge of the ungrammaticality of *Gapping-A (5–6 correct out of 6); knowledge of the contrast between Gapping-C and *Gapping-A (correct Gapping-C > correct *Gapping-A). The four criteria for the interpretation contrast were as follows: knowledge of the possibility of VPE with a subject reading (VPE-SR; 3–4 correct out of 4); knowledge of the possibility of Gapping with an object reading (Gapping-OR; 3–4 correct out of 4); knowledge of the contrast between VPE-SR and *VPE-OR (correct VPE-SR > correct *VPE-OR); knowledge of the impossibility of VPE with an object reading (*VPE-OR; 3–4 correct out of 4). Shading indicates participants who passed all four criteria for the grammaticality contrast and/or for the interpretation contrast.

Table 8.2

*Developmental Sequence of the Grammaticality and Interpretation Contrasts between VPE and Gapping: Early L2ers in Proficiency Order (*n *= 27)*

Proficiency		PPT	Grammaticality contrast				Interpretation contrast				Developmental stage
			Knowledge of the grammaticality of VPE	Knowledge of the ungrammaticality of *backward Gapping	Knowledge of the ungrammaticality of *Gapping-A	Knowledge of the contrast between Gapping-C and *Gapping-A	Knowledge of the possibility of VPE-SR	Knowledge of the possibility of Gapping-OR	Knowledge of the contrast between VPE-SR and *VPE-OR	Knowledge of the impossibility of *VPE-OR	
−6.65	(L)	EL2_11				✓	✓	✓			(1) Knowledge of the contrast between Gapping-C and *Gapping-A and the possibility of VPE-SR and Gapping-OR
−5.98	(L)	EL2_04		✓		✓	✓	✓			
−3.80	(L)	EL2_23	✓			✓	✓	✓	✓		
−3.41	(L)	EL2_12		✓	✓					✓	
−3.25	(L)	EL2_24				✓	✓				
−2.65	(L)	EL2_05				✓	✓	✓	✓		(2) Knowledge of the contrast between VPE-SR and *VPE-OR
−2.52	(L)	EL2_21	✓	✓	✓		✓	✓	✓	✓	
−2.49	(L)	EL2_10						✓			
−2.36	(L)	EL2_07			✓	✓				✓	
−2.35	(L)	EL2_08	✓				✓	✓	✓		(3) Knowledge of the grammaticality of VPE
−1.96	(L)	EL2_27	✓			✓	✓	✓	✓	✓	
−1.10	(M)	EL2_18				✓	✓	✓			
−0.78	(M)	EL2_25			✓	✓	✓	✓	✓		(4) Knowledge of the ungrammaticality of *Gapping-A
−0.71	(M)	EL2_15	✓	✓	✓	✓	✓	✓	✓	✓	(5) Knowledge of the ungrammaticality of *backward Gapping
−0.19	(M)	EL2_02	✓	✓	✓		✓	✓	✓		
0.39	(M)	EL2_03	✓		✓	✓		✓			
0.51	(M)	EL2_06	✓	✓	✓	✓	✓	✓	✓	✓	(6) Knowledge of the impossibility of *VPE-OR
0.51	(M)	EL2_26		✓		✓	✓	✓	✓	✓	
0.53	(M)	EL2_01	✓		✓	✓	✓	✓	✓		
0.80	(M)	EL2_22	✓	✓	✓	✓	✓	✓	✓	✓	
1.45	(H)	EL2_09			✓	✓	✓	✓	✓	✓	
1.62	(H)	EL2_13	✓		✓		✓	✓	✓	✓	
1.69	(H)	EL2_19	✓	✓	✓		✓	✓	✓		
1.90	(H)	EL2_17	✓	✓	✓	✓	✓	✓	✓	✓	
2.79	(H)	EL2_20	✓	✓	✓		✓	✓	✓	✓	
2.90	(H)	EL2_16	✓	✓	✓		✓	✓	✓	✓	
3.51	(H)	EL2_14	✓			✓	✓	✓	✓	✓	

Notes. (L), (M), and (H) respectively indicate the Lower proficiency group, the Medium proficiency group, and the Higher proficiency group. The check mark (✓) indicates presence of knowledge. The four criteria for the grammaticality contrast were as follows: knowledge of the grammaticality of VPE (8–12 correct out of 12); knowledge of the ungrammaticality of *backward Gapping (3 correct out of 3); knowledge of the ungrammaticality of *Gapping-A (5–6 correct out of 6); knowledge of the contrast between Gapping-C and *Gapping-A (correct Gapping-C > correct *Gapping-A). The four criteria for the interpretation contrast were as follows: knowledge of the possibility of VPE with a subject reading (VPE-SR; 3–4 correct out of 4); knowledge of the possibility of Gapping with an object reading (Gapping-OR; 3–4 correct out of 4); knowledge of the contrast between VPE-SR and *VPE-OR (correct VPE-SR > correct *VPE-OR); knowledge of the impossibility of VPE with an object reading (*VPE-OR; 3–4 correct out of 4). Shading indicates participants who passed all four criteria for the grammaticality contrast and/or for the interpretation contrast.

Table 8.3

*Developmental Sequence of the Grammaticality and Interpretation Contrasts between VPE and Gapping: Late L2ers in Proficiency Order (*n = *30)*

Proficiency		PPT	Grammaticality contrast				Interpretation contrast				Developmental stage
			Knowledge of the grammaticality of VPE	Knowledge of the ungrammaticality of *backward Gapping	Knowledge of the ungrammaticality of *Gapping-A	Knowledge of the contrast between Gapping-C and *Gapping-A	Knowledge of the possibility of VPE-SR	Knowledge of the possibility of Gapping-OR	Knowledge of the contrast between VPE-SR and *VPE-OR	Knowledge of the impossibility of *VPE-OR	
−4.36	(L)	LL2_26		✓	✓	✓	✓	✓	✓	✓	N/A
−2.31	(L)	LL2_22	✓	✓	✓	✓	✓	✓	✓	✓	
−2.00	(L)	LL2_05	✓	✓		✓	✓	✓	✓	✓	
−1.36	(L)	LL2_24	✓				✓	✓	✓	✓	
−0.79	(M)	LL2_18	✓	✓	✓	✓	✓	✓	✓	✓	
−0.31	(M)	LL2_08	✓		✓		✓	✓	✓	✓	
−0.24	(M)	LL2_19	✓	✓	✓	✓	✓	✓	✓	✓	
0.14	(M)	LL2_01	✓	✓	✓	✓	✓	✓	✓		
0.14	(M)	LL2_20	✓	✓	✓	✓	✓	✓	✓	✓	
0.22	(M)	LL2_16	✓	✓	✓	✓	✓	✓	✓	✓	
0.24	(M)	LL2_14	✓	✓	✓	✓	✓	✓	✓	✓	
0.25	(M)	LL2_04	✓		✓		✓		✓	✓	
0.46	(M)	LL2_11	✓	✓	✓	✓	✓	✓	✓	✓	
0.59	(M)	LL2_21	✓		✓		✓	✓	✓	✓	
0.64	(M)	LL2_10	✓				✓	✓	✓	✓	
0.80	(M)	LL2_27	✓	✓	✓	✓	✓	✓	✓	✓	
0.72	(M)	LL2_28	✓	✓	✓	✓	✓	✓	✓	✓	
0.90	(M)	LL2_02	✓				✓	✓	✓	✓	
1.44	(H)	LL2_23	✓		✓		✓	✓	✓	✓	
1.44	(H)	LL2_06	✓	✓	✓	✓	✓	✓	✓	✓	
1.52	(H)	LL2_13	✓	✓	✓	✓	✓	✓	✓	✓	
1.85	(H)	LL2_17	✓	✓	✓	✓	✓	✓	✓	✓	
2.09	(H)	LL2_07	✓	✓	✓	✓	✓	✓	✓	✓	
2.16	(H)	LL2_15	✓	✓	✓	✓	✓	✓	✓	✓	
2.39	(H)	LL2_09	✓	✓	✓	✓	✓	✓	✓	✓	
2.39	(H)	LL2_29	✓	✓	✓	✓	✓	✓	✓	✓	
2.41	(H)	LL2_25	✓	✓	✓	✓	✓		✓	✓	
2.80	(H)	LL2_03	✓				✓	✓	✓	✓	
3.53	(H)	LL2_12	✓	✓	✓	✓	✓		✓	✓	
3.84	(H)	LL2_30	✓	✓	✓	✓	✓	✓	✓	✓	

Notes. (L), (M), and (H) respectively indicate the Lower proficiency group, the Medium proficiency group, and the Higher proficiency group. The check mark (✓) indicates presence of knowledge. The four criteria for the grammaticality contrast were as follows: knowledge of the grammaticality of VPE (8–12 correct out of 12); knowledge of the ungrammaticality of *backward Gapping (3 correct out of 3); knowledge of the ungrammaticality of *Gapping-A (5–6 correct out of 6); knowledge of the contrast between Gapping-C and *Gapping-A (correct Gapping-C > correct *Gapping-A). The four criteria for the interpretation contrast were as follows: knowledge of the possibility of VPE with a subject reading (VPE-SR; 3–4 correct out of 4); knowledge of the possibility of Gapping with an object reading (Gapping-OR; 3–4 correct out of 4); knowledge of the contrast between VPE-SR and *VPE-OR (correct VPE-SR > correct *VPE-OR); knowledge of the impossibility of VPE with an object reading (*VPE-OR; 3–4 correct out of 4). Shading indicates participants who passed all four criteria for the grammaticality contrast and/or for the interpretation contrast.

The L1 children's data indicated five distinct developmental stages. First, at age 5;4, they concurrently acquired knowledge of (a) the grammaticality of VPE and (b) the possibility of VPE with a subject reading and Gapping with an object reading. Then they came to know the interpretation contrast between VPE with a subject reading and VPE with an object reading at age 5;6. The next stage, in which children showed they know that backward Gapping is ungrammatical and that VPE with an object reading is impossible, emerged by age 5;10. Then they showed knowledge of the ungrammaticality of Gapping in adjunct clauses at age 6;3. Lastly, at age 6;6, knowledge of the grammaticality contrast between Gapping in conjunct clauses and Gapping in adjunct clauses developed.

Regarding the early L2ers' data, six developmental stages were observed. The first stage was knowledge of (a) the grammaticality contrast between Gapping in conjunct clauses and Gapping in adjunct clauses and (b) the possibility of VPE with a subject reading and Gapping with an object reading. In the second stage, knowledge of the interpretation contrast between VPE with a subject reading and VPE with an object reading developed. Subsequently, the early L2ers showed they know that VPE (in conjunct and adjunct clauses) is grammatical. Next came knowledge of the ungrammaticality of Gapping in adjunct clauses, which was followed by knowledge of the ungrammaticality of backward Gapping. Finally, knowledge of the impossibility of VPE with an object reading emerged.

Because the majority of the late L2ers passed all the criteria, it was not possible to identify developmental sequences in the data from these participants.

Importantly, the individual analyses revealed *Age* effects in the L1 child data and *Proficiency* effects in the early L2er data; more learners displayed the target contrasts as *Age* increased in the case of the L1 children and as *Proficiency* increased in the case of the early L2ers. This analysis also found that one L1 child (aged 6;6) out of 24, five early L2ers out of 27, and 17 late L2ers out of 30 manifested robust knowledge of both the grammaticality contrast and the interpretation contrast between VPE and Gapping.

8.3 Theoretical Implications

8.3.1 Learnability.

8.3.1.1 Grammaticality contrast between VP-Ellipsis and Gapping.

Despite the fact that the grammaticality contrast between VPE and Gapping presents L1-English children and L1-Korean L2ers of English with learnability challenges, the older children and many of the higher-proficiency L2ers tested in this dissertation were able to develop implicit knowledge of this contrast. How did they come to have the contrast at issue, and what were the mechanisms that led to development beyond where they started and, at least for some, to converge on the target grammar? One possible answer to these questions comes from the learning-by-parsing approach to language acquisition (for L1 acquisition, see Berwick & Weinberg, 1984; Carroll, 2001; Fodor, 1998; Fodor & Sakas, 2017; Gibson & Wexler, 1994; Sakas & Fodor, 2012; Westergaard, 2014; for L2 acquisition, see Dekydtspotter & Renaud, 2014; Schwartz & Sprouse, 1994, 1996; Sharwood Smith & Truscott, 2014; Slabakova, 2016; Sprouse, 2011; Truscott & Sharwood Smith, 2004; Westergaard, 2019; White, 1987, 2003). This approach assumes that learners possess innate knowledge of the hypothesis space in which natural language grammars occur, i.e., Universal Grammar (UG), and do their best to parse the sentences they encounter using their existing grammars. Learning occurs when those grammars fail to provide an adequate parse for the sentence and learners are therefore forced to search through the store of parses that UG makes available for a viable alternative. If an attempted parse brings success, it will be integrated into the learner's grammar; if it repeatedly leads to failure, it will eventually be excluded from the grammar.

The main challenge that the L1-Engish children faced in overcoming the learnability problems at issue was figuring out which sentences are ungrammatical (see §2.4). One thing the L1 children needed to know is that Gapping is ungrammatical in adjunct clauses. However, the L1 children tested in this dissertation did not consistently reject Gapping in adjunct clauses until age 6;3. (The age of the youngest child who reliably rejected Gapping in adjunct clauses was 4;0.) I speculate that some of the children younger than 6;3 did not show target-like performance because they did not know that Gapping involves Across-the-Board (ATB) movement and that ATB movement is restricted to coordination.[1] Because Gapping does not display any detectable

[1] I thank Shin Fukuda (personal communication, 5 February 2020) for helpful discussion regarding this speculation.

movement of a constituent (unlike, say, the *wh*-fillers in English sentences containing filler–gap dependencies), children are likely to have difficulty figuring out that Gapping involves movement. If this idea is correct, L1 children should show target-like judgments for ATB movement sentences with a *wh*-filler—e.g., correctly accepting (1a) and rejecting (1b)—before being able to judge ATB movement sentences without a *wh*-filler (e.g., (1c) and (1d)) in a target-like manner.

(1) a. Sam likes pizza which Mom ate and Dad made.
(ATB movement with a *wh*-filler; conjunct)

b. *Sam likes pizza which Mom ate because Dad made.
(ATB movement with a *wh*-filler; adjunct)

c. Sam thinks that Mom ate pizza and Dad pasta.
(ATB movement without a *wh*-filler; conjunct)

d. *Sam thinks that Mom ate pizza because Dad pasta.
(ATB movement without a *wh*-filler; adjunct)

Furthermore, the way the AJT was implemented was probably not child-friendly enough, and this is likely to have affected the results. Admittedly, the AJT had rather many sentences to judge ($k = 68$), especially for 3- to 7-year-old children. These sentences were also presented without any context, which might have made it harder for children to process them. The fact that a written sentence was presented along with its corresponding audio stimulus might also have imposed a cognitive burden on immature readers; by the same token, pre-literate children had no written support while older, proficient readers did.[2] Future work should conduct AJTs that are more child-friendly by: (a) reducing the number of total sentences to judge, (b) having a brief context sentence precede each target sentence, and (c) presenting experimental sentences solely in an oral format and repeating them as often as the child wants. Incorporating such changes will, I believe, reveal convergence on the target grammar in younger children.

Another thing that the children needed to learn in order to acquire the target grammaticality contrast is that English does not allow backward Gapping. Interestingly, the

[2] My thanks to Kamil Deen (personal communication, 15 November 2019) and Bonnie D. Schwartz (personal communication, 7 April 2020) for helpful discussions about this issue.

children younger than age 5;10 who I tested sometimes did not reject sentences with backward Gapping. As was the case with Gapping in adjunct clauses, it is possible that the younger children did not know that Gapping involves ATB movement. Again, the design of the AJT was not as child-friendly as it could have been, which should be addressed in further research.

Despite the learnability challenges that young children face, the data reported in this dissertation demonstrate that most children aged 5;11 and older do in fact display implicit knowledge that English prohibits Gapping in adjunct clauses (see Table 8.1). It is possible that the older children performed better than the younger ones because they had received more relevant input, which led them to develop a grammar that allows ATB movement. Specifically, sentences containing a moved *wh*-filler, as in (1a) and (2), could provide a cue for ATB movement.

(2) Which book$_i$ does [Peter like t_i] and [Susan hate t_i]?

(de Vries, 2017, p. 1, (1))

As for backward Gapping, children's rejection of this pattern indicates that they indeed observe the constraint on Gapping direction (O'Grady, 1999), which states that a language's head-complement order predicts "the impossibility of a particular gapping direction without implying that the reverse gapping direction is permitted" (p. 143; see also §2.2).

The L1-Korean L2ers of English faced a different task in acquiring the target grammaticality contrast between VPE and Gapping. First, they needed to learn that VPE is possible in adjunct clauses. The reason for this is that Korean does not have VPE and even the closest analogues to VPE in Korean, i.e., Argument Ellipsis (AE), *Kulay* 'Do So' Anaphora, and Pseudo-VPE, cannot occur in adjunct clauses except for two cases: AE is allowed in adjunct clauses when an overt argument in the ellipsis clause occurs with the nominative case marker (vs. *-to* 'also'); *Kulay* 'Do So' Anaphora is permitted in adjunct clauses only when its antecedent precedes it in a separate sentence or main clause (see §2.1). Study 2 found that both the early L2ers and the late L2ers had developed implicit knowledge that VPE is grammatical in adjunct clauses (see Chapter 5).

Given that VPE hardly ever occurs in adjunct clauses in the input (Chapter 4), encountering sentences with another, higher frequency type of deletion in adjunct clauses might

serve as the needed trigger for acquiring implicit knowledge that VPE in adjunct clauses is grammatical. For example, cases of Nominal Ellipsis[3] in adjunct clauses (Günther, 2012; Lobeck, 1995), as in (3), may provide indirect evidence that other types of ellipsis can also occur in adjunct clauses in English.

(3) a. Because [$_{DP}$ all/both [e]] are so popular, these wines will probably be very expensive.
(Lobeck, 1995, p. 74, (10c))

b. Sally discussed the committee's decision to vote yes before she mentioned [$_{DP}$ the president's [e]].
(Lobeck, 2008, p. 156, (44c))

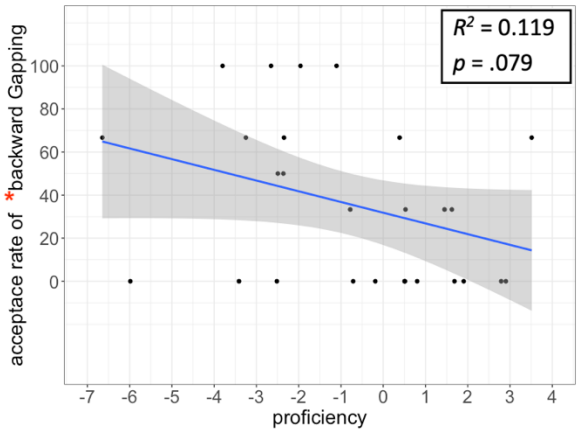
To test for this possibility, I analyzed a subset of data that I used in Study 1 (Chapter 4), i.e., the spoken input consisting of 2,144 utterances from a high school EFL textbook. While this input did not reveal a single instance of VPE in an adjunct clause, it did have two instances of Nominal Ellipsis (0.09%; e.g., *We have a variety of scarves here. … Well, if you buy two, you get a 10% discount*.). Admittedly, the frequency of Nominal Ellipsis in adjunct clauses is also low, albeit higher than the incidence of VPE in adjunct clauses. This possibility is nonetheless, I believe, worth exploring.

As for Gapping, since Korean permits only backward Gapping, the L2ers also needed to restructure their Interlanguage grammars so as to prohibit backward Gapping in English. Assuming that the L2ers already know that Gapping involves ATB movement from their L1 grammar, what they needed to learn is the direction of Gapping, which observes O'Grady's (1999) constraint. In the data from the early L2ers in this dissertation, a clear proficiency effect was observed such that lower-proficiency learners tended to accept backward Gapping and higher-proficiency learners tended to reject it, as shown in Figure 8.1 (for the individual results, see also Table 8.2).

[3] The analysis of Nominal Ellipsis is the same as that of VPE in that the head licenses the deletion of its complement. In the case of Nominal Ellipsis, the deletion of the NP is licensed by the D head (Chao, 1987; see also Lobeck, 1995).

Figure 8.3

*Relation between Proficiency and Acceptance Rate of *Backward Gapping for the Early L2ers in Study 2*



Note. The shaded region shows the 95% confidence interval for the sample mean.

A simple regression analysis performed on the judgment data from the early L2ers found that there was a marginally significant effect of *Proficiency* on the acceptance rates for backward Gapping ($\beta = -0.050$, $SE = 0.027$, $p = .079$). This result is in line with the idea that the initial state of L2 acquisition is the L1 grammar and that restructuring takes place as L2 proficiency increases, thus supporting the Full Transfer/Full Access model (Schwartz & Sprouse, 1996).

A regression analysis performed on the data from the late L2ers, by contrast, did not find a significant effect of *Proficiency* on the acceptance rates for backward Gapping ($\beta = 0.004$, $SE = 0.032$, $p = .903$). I concluded that the late L2ers had likely already developed the target grammaticality contrast between VPE and Gapping (see §5.4.2), perhaps because they had received more chances to test out their parses than had the early L2ers; the late L2ers had higher proficiency scores (see §5.2.3) and more years of exposure to English than the early L2ers did (late L2ers: $M = 14.20$, $SD = 2.93$; early L2ers: $M = 3.56$, $SD = 1.67$).

Note that in the by-group analyses, neither the early L2ers nor the late L2ers showed target-like performance on forward Gapping (i.e., Gapping in conjunct clauses), presumably due, I argued, to the processing difficulty involved in this sentence type (see §5.4.1). Nevertheless, a further individual analysis (see §5.4.2) revealed that six of the late L2ers displayed robust knowledge of forward Gapping, accepting 5 or more out of 6 items in that condition; these L2ers were also able to correctly reject all 3 instances of the ungrammatical backward Gapping pattern. I believe that an AJT design that adds a context sentence before each target sentence or uses

stimuli with a PP following the verb gap (e.g., *Mom sat on the sofa, and Dad [e] on the chair.*) could help reduce participants' processing load, thereby facilitating the acceptance of forward Gapping. I leave this issue for further investigations.

8.3.1.2 Interpretation contrast between VP-Ellipsis and Gapping.

We turn now to the interpretation contrast between VPE and Gapping. L1 children needed to create a grammar that rules out VPE with an object reading (VPE-OR). Study 3 found that the children came to have this knowledge as early as age 5;6. Surprisingly, some of the L1 children did not fully disallow VPE-OR. It might be tempting to conclude that these children did not know that VPE-OR is impossible in English, but this seems very unlikely given that previous research has found that children as young as age 3 show robust knowledge about which interpretations of elided pronouns in VPE are grammatical and ungrammatical (Foley et al., 2003) and that children as young as 3;11 fully master the parallelism effects associated with VPE (Matsuo & Duffield, 2001).

A more promising alternative is that the children sometimes mistakenly interpreted VPE sentences, as in (4), as Stripping sentences, as in (5) (Johnson, 2019; Repp, 2009; Winkler, 2005); these sentences can be ambiguous, allowing the argument following the conjunction (e.g., *Dad*) to be interpreted as either the subject or the object (Ruijgrok, 2018).[4, 5]

(4) Mom hugged the boy at home and Dad did [e] too.

[4] Stripping elides all non-contrastive material in the second conjunct clause, leaving a single contrasting remnant to be focused (Merchant, 2003, 2004), as illustrated in (i). The remnant is accompanied by a polarity element, such as *not* or *too*. (ia) and (ib) provide analyses for the subject and object readings of Stripping, respectively.

(i) Mom hugged the boy and [Dad [e] too[a]]. / and [[e] Dad too[b]].
   a. Subject reading: [[$_{CP}$ [$_{TP}$ Mom hugged the boy] and [$_{CP}$ [$_{FocP}$ [$_{NP}$ Dad$_i$] [$_{TP}$ ~~t$_i$ hugged the boy~~ too.]]]]]
   b. Object reading: [[$_{CP}$ [$_{TP}$ Mom hugged the boy] and [$_{CP}$ [$_{FocP}$ [$_{NP}$ Dad$_i$] [$_{TP}$ ~~Mom hugged t$_i$~~ too.]]]]]

[5] There is no L1 acquisition research on Stripping so far. My pilot study conducted in Spring 2016 showed that four mature English native speakers all preferred the object reading over the subject reading for stripping sentences (e.g., *John loves his mother and Peter too.*).

(5) Mom hugged the boy at home and [Dad [e] too[a]]. / [[e] Dad too[b]].

a. Subject reading: 'Mom hugged the boy at home, and Dad (hugged the boy at home) too.'

b. Object reading: 'Mom hugged the boy at home and (Mom hugged) Dad (at home) too.'

If this is what the children were doing, then their acceptance of VPE-OR trials would indicate that they were unable to reanalyze the sentence even after encountering the auxiliary verb *did* (for reanalysis problems in children, see Trueswell, Sekerina, Hill & Logrip, 1999; see also Kidd, Stewart & Serratrice, 2011). But this raises the question of why the younger L1 children in this study experienced such reanalysis problems while those in previous studies on the interpretation of elided pronouns (e.g., Foley et al., 2003) and parallelism effects (e.g., Matsuo & Duffield, 2001) did not, instead displaying early mastery of VPE.

This difference can be explained in terms of similarity-based interference (Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006; Villata et al., 2018). The target VPE sentences in Study 3 always contained three arguments (e.g., *Mom*, *the boy*, *Dad*), all of which were animate. The fact that all three arguments shared the same [+ animate] feature might have increased the memory load for (i.e., induced similarity-based interference in), especially, the child participants, thereby making it more difficult for them to parse the ellipsis clause accurately. The VPE sentences used in previous L1 studies in this line of research, by contrast, did not suffer from this problem. In Foley et al.'s (2003) study, for example, whereas the subjects of both the antecedent clause and the ellipsis clause were always [+ animate], the object in the antecedent clause was always [– animate] (e.g., *Oscar bites his apple and Bert does too.*), and this animacy mismatch might have made it easier for children to parse the ellipsis clauses correctly.

Future research should investigate the possibility that reanalysis problems, which are known to characterize processing in young L1 children, influence children's interpretations of sentences with VPE. It would be worth testing whether repeating the stimulus sentence facilitates the hypothesized reanalysis from Stripping to VPE (for evidence that repetition helps young children reanalyze passive sentences, see Deen et al., 2018) or whether making the prosody of the auxiliary verb *did* more prominent by lengthening its duration or widening its pitch range helps promote this hypothesized reanalysis (for evidence that manipulation of prosody helps young children override their preferred interpretation for sentences with ambiguous

PP-attachment, see Diehl, Friedberg, Paul, & Snedeker, 2014), thus leading younger children to perform in a more target-like manner.

The L1-Korean L2ers of English also needed to come to rule out VPE-OR in order to develop target-like knowledge of the interpretation contrast at issue. This is because Korean does not have VPE, and all of the closest analogues of VPE in Korean (i.e., Argument Ellipsis, *Kulay* 'Do So' Anaphora, Pseudo-VPE) do allow the object reading. I have proposed that L2ers attempt to parse the input that they encounter on the basis of the current grammar, and that restructuring of the Interlanguage grammar occurs when the existing grammar consistently fails to provide adequate parses for sentences in the input. While the lower-proficiency members of the early L2er group tended to allow VPE-OR, the higher-proficiency ones consistently disallowed it (see Figure 6.5). Again, this result suggests that the initial state of the Interlanguage grammar is the L1 grammar and that L2ers eventually converge on the target grammar as their proficiency increases, in line with the Full Transfer/Full Access model (Schwartz & Sprouse, 1996). The absence of proficiency effects in the data from the late L2ers reflects the fact that most of them already had knowledge of the target interpretation contrast, presumably due to their high proficiency (see §5.2.3) and longer duration of exposure to the TL (see §8.3.1.1).

8.3.2 Fundamental identity between L1 acquisition and L2 acquisition.

The Fundamental Difference Hypothesis (FDH; Bley-Vroman, 1989, 1990) maintains that whereas children have the principles and parameters of UG at their disposal during L1 acquisition, adults learning an L2 only have access to the L1 grammar and general problem-solving strategies. According to Song and Schwartz (2009), this hypothesis can be tested in two ways. The first way is to investigate whether adults can overcome L2 learnability problems, i.e., cases where knowledge of the target phenomenon cannot be attributed to the L1, Target Language input, or instruction (Schwartz & Sprouse, 2000). The second way is to compare the developmental sequences of adult and child L2ers with the same L1; the presence of similar developmental trajectories in L2 adults and children would indicate that the same mechanisms guide the acquisition process for both groups and therefore that UG is involved not only in child L2 acquisition but also in adult L2 acquisition (Schwartz 1987, 1992, 2004).

There is no clear-cut boundary between adult and child L2ers, but adult L2ers have often been defined as those who begin to learn the TL after the critical period for language acquisition

has ended, at which point target-like language attainment is claimed to be no longer possible due to physiological changes in the brain (Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009; DeKeyser & Larson-Hall, 2005; Hyltenstam & Abrahamsson, 2003; Lenneberg, 1967; Long, 2005). Because Haznedar (2013) places the end of the critical period at age 7–10, I decided to operationalize adult L2ers as those whose age of onset was greater than 10.

However, although Studies 2 and 3 targeted phenomena that are thought to constitute learnability problems, thus making it possible to test the FDH, I did not specifically recruit adult L2ers to participate in these studies. In fact, only two of the late L2ers had started learning English after the age of 10: LL2_17 (age of onset: 11) and LL2_20 (age of onset: 12). I was not able to identify any developmental sequence these two adult L2ers passed through because they met all the criteria for displaying target-like knowledge of the grammaticality and interpretation contrasts between VPE and Gapping (see Table 8.3). Furthermore, I could not make comparisons between the early L2ers and the late L2ers in terms of the developmental sequences, either, since these two groups were not comparable in terms of proficiency. The late L2ers, most of whom showed the target grammaticality and interpretation contrasts between VPE and Gapping, had relatively high proficiency (see §5.2.3). Nevertheless, the studies clearly showed that with respect to the phenomena under investigation, the two adult L2ers identified above had converged on a grammar that was ostensibly identical to that of the native speakers, the older L1 children, and the higher-proficiency early L2ers despite the learnability problems they faced. This result suggests that UG is indeed involved in adult L2 acquisition, thus providing evidence against the FDH. Indeed, the result is more in line with the Fundamental Identity Hypothesis (Hopp, 2007), which maintains that grammatical representations can in principle come to be essentially identical in L1 acquisition and (child and adult) L2 acquisition.

8.3.3 Fundamental identity between L1 processing and L2 processing.

There are two main theoretical approaches to understanding how L2 sentence processing relates to L1 sentence processing. One maintains that even very advanced L2ers differ from native speakers when processing sentences (Boxwell & Felser, 2017; Felser, 2019; Felser & Roberts, 2007; Felser, Roberts, Marinis, & Gross, 2003; Jackson & van Hell, 2011; Jiang, Hu, Chrabaszcz, & Ye, 2017; Juffs & Rodriguez, 2014; Marinis, Roberts, Felser, & Clahsen, 2005; Papadopoulou & Clahsen, 2003; Zheng & Lemhöfer, 2019) because the parsing mechanisms

involved in the L1 and L2 are fundamentally different. For example, the Shallow Structure Hypothesis (Clahsen & Felser, 2006a, 2006b; see also Clahsen & Felser, 2018) holds that in real-time processing, L2ers are not able to build fully elaborated syntactic representations and so they instead rely heavily on lexical and pragmatic information. The other approach posits that L2ers can use detailed morphosyntactic information and that they are therefore able to come to process sentences in a target-like manner. According to this view, L1 and L2 processing are (or at least can become) qualitatively the same (Alemán Bañón, Fiorentino, & Gabriele, 2014; Cunnings, 2017; Fernandez, Höhle, Brock, & Nickels, 2018; Herbay, Gonnerman, & Baum, 2018; Hopp, 2007, 2018) and any differences that are observed arise (primarily) from inefficient lexical access (e.g., Dekydtspotter, Schwartz, & Sprouse, 2006; Hopp, 2007, 2018; McDonald, 2006), reduced working memory capacity (e.g., Kaan, Ballantyne, & Wijnen, 2015), or low proficiency (e.g., Cunnings, 2017; Omaki & Schulz, 2011; Wen, Miyao, Takeda, Chu, & Schwartz, 2010; Witzel, Witzel, & Nicol, 2012) on the part of the L2ers.

L1-Korean L2ers' processing of English Gapping provides an ideal test case for investigating the nature of L2 processing mechanisms. For one thing, Gapping is a complex and rare syntactic phenomenon, which makes it very likely that L2ers will have difficulty using, in real time, syntactic information to process sentences containing Gapping. Also, Gapping sentences do not provide an overt signal that movement has occurred, unlike the *wh*-fillers that appear in *wh*-questions and relative clauses. Furthermore, it is reasonable to assume that L1-Korean L2ers might have trouble resolving verb gaps in English due to cross-linguistic differences in the direction of Gapping: Because the verb gap precedes its antecedent in Korean, the dependency in Korean can be resolved only after encountering the overt verb in the non-gapped clause. In English, on the other hand, the verb gap follows its antecedent, thus making it necessary to resolve the dependency at the gap site. Adjusting to this difference may be difficult for L1-Korean L2ers of English, especially during real-time sentence processing.

Study 4 used a self-paced reading task that exploited the fact that plausibility effects can be induced in Gapping sentences by changing the verb (e.g., *Bill {ordered/*drank} coffee and tea at the cafe, and Jane sandwiches and cake at the bakery.*). For a baseline, the task included VPE counterparts (e.g., *Bill {ordered/drank} coffee and tea at the cafe, and Jane did too at the bakery.*) as critical sentences and *where*-clause counterparts (e.g., *Bill {ordered/drank} coffee and tea at the cafe where sandwiches and cake were very popular*.) as fillers since they do not

exhibit plausibility effects. Despite the challenges that they faced, the adult L2ers performed just like the native speakers in displaying plausibility effects for only the Gapping sentences, which indicates that they were successful at identifying verb gaps and reconstructing verbal information at the gap site. Note that this result cannot have come from some sort of shallow structure processing, such as non-syntactic surface word association procedures (for discussion, see §7.6). Instead, it could only have come from the use of detailed syntactic information during real-time sentence processing. This finding therefore suggests that L2ers' parsing mechanisms are not qualitatively different from those of native speakers.

8.4 Conclusion

To conclude: Target-like knowledge of the grammaticality and interpretation contrasts between VPE and Gapping in English was shown by the L1-English children (by the age of 5;11 for the grammaticality contrast and by the age of 5;6 for the interpretation contrast) and by the early and late L1-Korean L2ers of English with higher proficiency. These findings indicate that the participants had overcome the learnability problems they faced, thereby providing evidence for the operation of UG in both L1 and L2 acquisition. Furthermore, the target-like processing patterns observed in the adult L2ers indicate that the parsing mechanisms employed by native speakers and L2ers are qualitatively the same.

Appendix A:
Proficiency Data from the Picture Narration Task

As an independent measure of English proficiency in Study 2 (Chapter 5) and Study 3 (Chapter 6), a picture narration task (PNT; K.-S. Park, 2014; Unsworth, 2005; Whong-Barr & Schwartz, 2002) was administered to all 43 early L2ers and all 31 late L2ers (and only 32 of the 70 L1 adults and 32 of the 46 L1 children). Recall that 13 L1 children, 14 early L2ers, and one late L2er were excluded from the screening procedure (see §5.2.1). Thus, only data from the remaining participants who completed the PNT (32 out of 70 L1 adults, 24 out of the 33 L1 children, 27 early L2ers, 30 late L2ers) were analyzed for proficiency.

In the PNT, participants were presented with 3 sets of 4 pictures depicting a series of events (see K.-S. Park, 2014, p. 146), and then were asked to tell a short story based on the 4 pictures that they saw in each set. I recorded participants' utterances using Praat (Boersma & Weenink, 2017). The recorded data were transcribed by a native speaker of English and me. I subsequently removed all filler words, self-corrected words, and repeated words from both of the transcripts. For example, for the utterance *Well, I know she's brushing teeth because she has, he has a maroon thing* from participant L1C_10[1], I removed the filler *well* and the self-corrected words *she has*. And for EL2_02's utterance *The book…book ripped into two pieces*, I removed the repeated word *book*. Next, the utterances were converted into lists of T-units for use in computing the participant's level of English proficiency. At this point in the process, seven utterances produced by two L2ers (EL2_05, LL2_26) were removed because they contained Korean words. All disparities between the two transcripts were resolved through discussion.

Following K.-S. Park (2014) and Unsworth (2005) for the computation of English proficiency scores, I used three measures to analyze the elicited production data from the PNT: (a) morpho-syntactic complexity, (b) lexical complexity, and (c) morphological/syntactic/lexical accuracy.

Morpho-syntactic complexity was measured in terms of verbal density by dividing the number of finite verbs plus the number of nonfinite verbs (infinitives, gerunds, and participles) by the total number of T-units (see K.-S. Park, 2014, p. 157). For lexical complexity, the Moving

[1] A unique participant identifier was assigned to each participant to ensure anonymity. For example, the identifiers L1A_01, L1C_01, EL2_01, and LL2_01 were given to the first participant tested and included for analysis in, respectively, the L1 adult group, L1 child group, early L2er group, and late L2er group.

Average Type-Token Ratio (MATTR; Covington & McFall, 2010) was computed by calculating the average of the type-token ratio for every moving text sequence of 15 consecutive words. I chose MATTR rather than Guiraud's index[2] (Guiraud, 1954)—which was used in the studies by K.-S. Park and Unsworth—because MATTR is less sensitive to text length effects and is therefore more suitable for comparing speech samples of different lengths (Zenker & Kyle, 2019; see also Covington & McFall, 2010).

Lastly, the rate of error-free T-units was the measure of morphological/syntactic/lexical accuracy. Morphological, syntactic, and lexical errors were manually coded by two English native speakers. For example, morphological errors included errors in subject–verb agreement (e.g., [EL2_04] *The girl brush the teeth.*), tense agreement (e.g., [EL2_06] *The bear woke up and say ...*), and adjectives (e.g., [LL2_10] *sleep* for *asleep*). Syntactic errors included errors in the use of overt determiners (e.g., [LL2_13] *So Ø boy argued that the book is too close to her.*) and voice (e.g., [LL2_01] *After that, the boy was waked and was afraid again.*). Lexical errors involved non-target-like use of target forms with respect to their meaning or function, such as non-target-like use of lexical items (e.g., [LL2_02] *... she visited her mom and dad to stay with her.*). For the full details about the error coding procedure, see K.-S. Park (2014, pp. 164–168). The interrater-reliability for error coding was high, with a Cohen's Kappa coefficient of 0.735 (see Landis & Koch, 1977). Points of disagreement that arose during the coding process were resolved through discussion between the coders and me.

Except for the accuracy coding, all the analysis of the production data from the PNT was done in Python. The python code is available at https://github.com/Haerim-Hwang/NLP_Python/tree/master/Measure_English_proficiency.

In completing the PNT, the L1 adults (n = 32), L1 children (n = 24), early L2ers (n = 27), and late L2ers (n = 30) produced a total of, respectively, 437 T-units, 333 T-units, 333 T-units, and 548 T-units. On average, the L1 adults produced 13.66 T-units (SD = 3.65), the L1 children 13.88 T-units (SD = 2.54), the early L2ers 12.33 T-units (SD = 2.79), and the late L2ers 18.27 T-units (SD = 11.24). The participants' verbal density scores, MATTR scores, and rates of error-free T-units are provided in Table A-1.

[2] Guiraud's index is computed "by dividing the number of different lexical types by the square root of the total number of tokens ($V/\sqrt{N}$)" (K.-S. Park, 2014, p. 160).

Table A-1
Verbal Density, Moving Average Type-Token Ratio (MATTR), and Rate of Error-Free T-Units per Group

	Verbal Density	**MATTR**	**Rate of Error-Free T-units**
L1 Adults	1.85	0.87	0.98
(*n* = 32)	(*SD* = 0.39)	(*SD* = 0.05)	(*SD* = 0.05)
L1 Children	1.35	0.82	0.85
(*n* = 24)	(*SD* = 0.33)	(*SD* = 0.05)	(*SD* = 0.14)
Early L2ers	1.35	0.83	0.46
(*n* = 27)	(*SD* = 0.30)	(*SD* = 0.06)	(*SD* = 0.26)
Late L2ers	1.53	0.86	0.53
(*n* = 30)	(*SD* = 0.30)	(*SD* = 0.04)	(*SD* = 0.19)

Based on the assumption that each of the three measures contributes to L2 proficiency to an equal extent (K.-S. Park, 2014), the three sub-scores for each participant were converted into standardized *z*-scores and then added together to produce a combined proficiency score. The verbal density scores, MATTR scores, rates of error-free T-units, and combined proficiency scores for each of the four participant groups are provided in, respectively, Tables A-2–A-5.

Table A-2
Proficiency Data: L1 Adults (n = 32)

Participant	Number of T-units	Verbal density		Lexical diversity		Accuracy		Proficiency *z*-score
		Verbs/ T-units	*z*-score	MATTR	*z*-score	Error-free T-units	*z*-score	
L1A_39	14	1.57	−0.74	0.8588	−0.28	1.00	0.44	−0.59
L1A_40	10	1.80	−0.14	0.8017	−1.82	1.00	0.44	−1.52
L1A_41	25	1.88	0.07	0.8917	0.60	1.00	0.44	1.11
L1A_42	12	2.00	0.39	0.8804	0.30	1.00	0.44	1.12
L1A_43	16	2.13	0.72	0.8899	0.55	1.00	0.44	1.71
L1A_44	10	2.40	1.44	0.9200	1.36	1.00	0.44	3.25
L1A_45	10	2.00	0.39	0.8624	−0.19	1.00	0.44	0.64
L1A_46	12	1.50	−0.93	0.9261	1.53	1.00	0.44	1.04
L1A_47	8	2.63	2.03	0.8764	0.19	1.00	0.44	2.67
L1A_48	18	2.44	1.56	0.8723	0.08	1.00	0.44	2.08
L1A_49	13	1.62	−0.63	0.7549	−3.08	1.00	0.44	−3.26
L1A_50	11	2.00	0.39	0.8876	0.49	1.00	0.44	1.32
L1A_51	19	1.53	−0.86	0.9006	0.84	0.84	−3.04	−3.07
L1A_52	13	1.85	−0.02	0.8984	0.78	0.85	−2.95	−2.19
L1A_53	12	1.67	−0.49	0.8426	−0.72	1.00	0.44	−0.77
L1A_54	17	1.59	−0.70	0.8658	−0.10	0.94	−0.86	−1.65
L1A_55	10	1.70	−0.40	0.9089	1.06	1.00	0.44	1.10
L1A_56	17	1.71	−0.39	0.8819	0.34	1.00	0.44	0.39
L1A_57	16	1.88	0.06	0.8683	−0.03	1.00	0.44	0.47
L1A_58	16	2.19	0.88	0.8763	0.19	0.94	−0.94	0.13
L1A_59	8	1.88	0.06	0.8908	0.58	0.88	−2.32	−1.68
L1A_60	12	2.08	0.61	0.8184	−1.37	1.00	0.44	−0.32
L1A_61	12	2.67	2.14	0.8808	0.31	1.00	0.44	2.89
L1A_62	9	1.67	−0.49	0.8695	0.00	1.00	0.44	−0.05
L1A_63	12	1.17	−1.81	0.8478	−0.58	1.00	0.44	−1.95
L1A_64	13	1.00	−2.25	0.7732	−2.59	1.00	0.44	−4.40
L1A_65	17	1.29	−1.48	0.8942	0.67	1.00	0.44	−0.37
L1A_66	17	2.00	0.39	0.8792	0.27	1.00	0.44	1.09
L1A_67	15	1.53	−0.84	0.8887	0.52	1.00	0.44	0.12
L1A_68	15	2.20	0.91	0.8692	−0.01	1.00	0.44	1.35
L1A_69	12	2.08	0.61	0.8506	−0.50	0.92	−1.40	−1.30
L1A_70	16	1.69	−0.44	0.8923	0.62	1.00	0.44	0.62

Table A-3
Proficiency Data: L1 Children (n = 24)

Participant code	Number of T-units	Verbal density		Lexical diversity		Accuracy		Proficiency *z*-score
		Verbs/ T-units	*z*-score	MATTR	*z*-score	Error-free T-units	*z*-score	
L1C_10	22	1.27	−0.24	0.8977	1.51	0.68	−1.25	0.03
L1C_11	14	1.07	−0.86	0.8243	0.03	0.79	−0.47	−1.30
L1C_12	14	1.14	−0.64	0.8420	0.39	1.00	1.12	0.87
L1C_13	13	1.23	−0.37	0.8171	−0.11	0.92	0.55	0.07
L1C_14	12	1.25	−0.31	0.7333	−1.80	0.67	−1.36	−3.47
L1C_15	13	1.85	1.53	0.8057	−0.34	0.77	−0.60	0.59
L1C_16	13	1.38	0.11	0.8914	1.39	0.62	−1.74	−0.25
L1C_17	13	1.38	0.11	0.7798	−0.87	1.00	1.12	0.36
L1C_18	18	1.44	0.29	0.8704	0.96	0.94	0.71	1.96
L1C_19	14	1.36	0.02	0.8483	0.52	0.93	0.59	1.13
L1C_20	13	1.00	−1.08	0.8705	0.97	0.77	−0.60	−0.71
L1C_21	13	1.15	−0.61	0.7505	−1.46	1.00	1.12	−0.94
L1C_22	17	1.24	−0.35	0.8196	−0.06	0.94	0.68	0.26
L1C_23	9	2.00	2.00	0.7858	−0.74	0.78	−0.53	0.73
L1C_24	13	2.38	3.19	0.9017	1.60	0.92	0.55	5.33
L1C_25	16	1.38	0.08	0.8825	1.21	0.94	0.65	1.94
L1C_26	15	1.53	0.56	0.7951	−0.56	0.93	0.62	0.63
L1C_27	13	1.54	0.58	0.8630	0.81	1.00	1.12	2.51
L1C_28	14	1.00	−1.08	0.7682	−1.10	1.00	1.12	−1.06
L1C_29	14	1.14	−0.64	0.7371	−1.73	0.86	0.06	−2.31
L1C_30	15	1.40	0.15	0.8538	0.63	0.87	0.13	0.91
L1C_31	12	1.08	−0.82	0.8333	0.22	0.50	−2.60	−3.21
L1C_32	11	1.09	−0.80	0.7829	−0.80	0.82	−0.23	−1.84
L1C_33	12	1.08	−0.82	0.7899	−0.66	0.75	−0.74	−2.23

Table A-4
Proficiency Data: Early L2ers (n = 27)

Participant code	Number of T-units	Verbal density		Lexical diversity		Accuracy		Proficiency *z*-score
		Verbs/ T-units	*z*-score	MATTR	*z*-score	Error-free T-units	*z*-score	
EL2_01	14	1.43	−0.04	0.8757	0.57	0.50	0.00	0.53
EL2_02	15	1.07	−1.23	0.8774	0.60	0.60	0.44	−0.19
EL2_03	13	1.46	0.06	0.7798	−1.20	0.85	1.53	0.39
EL2_04	13	1.00	−1.45	0.7189	−2.33	0.00	−2.20	−5.98
EL2_05	13	1.08	−1.20	0.8302	−0.27	0.23	−1.19	−2.65
EL2_06	12	1.50	0.19	0.8619	0.31	0.50	0.00	0.51
EL2_07	10	1.40	−0.14	0.7961	−0.90	0.20	−1.32	−2.36
EL2_08	11	1.00	−1.45	0.8721	0.50	0.18	−1.40	−2.35
EL2_09	16	1.69	0.80	0.8500	0.09	0.63	0.55	1.45
EL2_10	11	1.09	−1.15	0.8048	−0.74	0.36	−0.60	−2.49
EL2_11	5	1.20	−0.79	0.6476	−3.65	0.00	−2.20	−6.65
EL2_12	12	1.33	−0.36	0.7392	−1.96	0.25	−1.10	−3.41
EL2_13	12	1.50	0.19	0.8628	0.33	0.75	1.10	1.62
EL2_14	11	2.00	1.83	0.8817	0.68	0.73	1.00	3.51
EL2_15	13	1.08	−1.20	0.8802	0.65	0.46	−0.17	−0.71
EL2_16	16	1.81	1.21	0.8765	0.58	0.75	1.10	2.90
EL2_17	16	1.31	−0.42	0.8813	0.67	0.88	1.66	1.90
EL2_18	9	1.22	−0.72	0.8373	−0.14	0.44	−0.24	−1.10
EL2_19	16	1.75	1.01	0.8963	0.95	0.44	−0.27	1.69
EL2_20	13	1.69	0.82	0.8872	0.78	0.77	1.19	2.79
EL2_21	12	1.08	−1.18	0.7723	−1.34	0.50	0.00	−2.52
EL2_22	10	1.60	0.52	0.8364	−0.16	0.60	0.44	0.80
EL2_23	16	1.00	−1.45	0.7325	−2.08	0.44	−0.27	−3.80
EL2_24	6	1.00	−1.45	0.8667	0.40	0.00	−2.20	−3.25
EL2_25	13	1.23	−0.69	0.7759	−1.28	0.77	1.19	−0.78
EL2_26	13	1.77	1.07	0.8239	−0.39	0.46	−0.17	0.51
EL2_27	12	1.08	−1.18	0.8821	0.69	0.17	−1.47	−1.96

Table A-5
Proficiency Data: Late L2ers (n = 30)

Participant code	Number of T-units	Verbal density		Lexical diversity		Accuracy		Proficiency *z*-score
		Verbs/ T-units	*z*-score	MATTR	*z*-score	Error-free T-units	*z*-score	
LL2_01	14	1.36	−0.28	0.8507	0.11	0.57	0.32	0.14
LL2_02	15	1.73	0.95	0.8498	0.09	0.47	−0.15	0.90
LL2_03	16	1.94	1.62	0.8937	0.90	0.56	0.28	2.80
LL2_04	27	1.37	−0.24	0.8580	0.24	0.56	0.25	0.25
LL2_05	12	1.08	−1.18	0.7805	−1.19	0.58	0.37	−2.00
LL2_06	16	1.63	0.60	0.8752	0.56	0.56	0.28	1.44
LL2_07	21	1.67	0.74	0.8671	0.41	0.71	0.95	2.09
LL2_08	13	1.08	−1.20	0.8286	−0.30	0.77	1.19	−0.31
LL2_09	24	1.83	1.28	0.8848	0.74	0.58	0.37	2.39
LL2_10	13	1.54	0.32	0.8897	0.83	0.38	−0.51	0.64
LL2_11	12	1.33	−0.36	0.8493	0.08	0.67	0.74	0.46
LL2_12	16	1.50	0.19	0.9363	1.69	0.88	1.66	3.53
LL2_13	28	1.86	1.36	0.8620	0.32	0.46	−0.16	1.52
LL2_14	13	1.23	−0.69	0.8492	0.08	0.69	0.85	0.24
LL2_15	16	1.75	1.01	0.8324	−0.23	0.81	1.38	2.16
LL2_16	14	1.29	−0.51	0.8333	−0.21	0.71	0.95	0.22
LL2_17	30	1.87	1.39	0.8854	0.75	0.43	−0.29	1.85
LL2_18	12	1.25	−0.63	0.8162	−0.53	0.58	0.37	−0.79
LL2_19	16	1.31	−0.42	0.8401	−0.09	0.56	0.28	−0.24
LL2_20	25	1.76	1.04	0.8769	0.59	0.16	−1.50	0.14
LL2_21	13	1.31	−0.44	0.8732	0.52	0.62	0.51	0.59
LL2_22	14	1.29	−0.51	0.8496	0.09	0.07	−1.89	−2.31
LL2_23	15	1.53	0.30	0.9146	1.29	0.47	−0.15	1.44
LL2_24	13	1.23	−0.69	0.8914	0.86	0.15	−1.52	−1.36
LL2_25	12	1.83	1.28	0.8860	0.76	0.58	0.37	2.41
LL2_26	6	1.00	−1.45	0.7273	−2.18	0.33	−0.73	−4.36
LL2_27	70	1.47	0.10	0.8626	0.33	0.59	0.38	0.80
LL2_28	19	1.95	1.66	0.8510	0.11	0.26	−1.04	0.72
LL2_29	21	1.76	1.05	0.9006	1.03	0.57	0.32	2.39
LL2_30	12	2.08	2.10	0.8992	1.00	0.67	0.74	3.84

Appendix B:
Stimuli for the Acceptability Judgment Task

Instructions:

This is a 'smiley face' choice task. In this task, you will decide whether sentences are good/possible or bad/impossible in English. First, you will see a sentence on the screen while hearing that same sentence twice on the headphones. After you hear the sentence for the second time, a 'smiley face' scale will pop up on the screen. It is at this point that you can press a number on the keyboard to rate the sentence on a 4-point "smiley face" scale with an additional "I don't know" option.

Critical items

(a) VPE-Conjunct, (b) VPE-Adjunct, (c) Gapping-Conjunct, (d) *Gapping-Adjunct

1. (a) Sara made pizza, and Kelly did too.
   (b) Sara made pizza because Kelly did.
   (c) Sara made pizza, and Kelly pasta.
   (d) *Sara made pizza because Kelly pasta.

2. (a) George prepared pictures, and Sally did too.
   (b) George prepared pictures because Sally did.
   (c) George prepared pictures, and Sally flowers.
   (d) *George prepared pictures because Sally flowers.

3. (a) David ordered tea, and Kris did too.
   (b) David ordered tea because Kris did.
   (c) David ordered tea, and Kris coffee.
   (d) *David ordered tea because Kris coffee.

4. (a) Tom used a knife, and Alice did too.
   (b) Tom used a knife because Alice did.
   (c) Tom used a knife, and Alice scissors.
   (d) *Tom used a knife because Alice scissors.

5. (a) John borrowed books, and Bill did too.
   (b) John borrowed books because Bill did.
   (c) John borrowed books, and Bill DVDs.
   (d) *John borrowed books because Bill DVDs.

6. (a) Andrew played soccer, and Paul did too.
   (b) Andrew played soccer because Paul did.
   (c) Andrew played soccer, and Paul baseball.
   (d) * Andrew played soccer because Paul baseball.

7. (a) Bob wrote a diary, and Nate did too.
   (b) Bob wrote a diary because Nate did.
   (c) Bob wrote a diary, and Nate an essay.
   (d) * Bob wrote a diary because Nate an essay.

8. (a) Brian bought shoes, and Kevin did too.
   (b) Brian bought shoes because Kevin did.
   (c) Brian bought shoes, and Kevin a bag.
   (d) * Brian bought shoes because Kevin a bag.

9. (a) Andy got a robot, and Helen did too.
   (b) Andy got a robot because Helen did.
   (c) Andy got a robot, and Helen a doll.
   (d) * Andy got a robot because Helen a doll.

10. (a) Fred sold a car, and Ryan did too.
    (b) Fred sold a car because Ryan did.
    (c) Fred sold a car, and Ryan a truck.
    (d) * Fred sold a car because Ryan a truck.

11. (a) Ann watched a movie, and Tony did too.
    (b) Ann watched a movie because Tony did.
    (c) Ann watched a movie, and Tony a drama.
    (d) * Ann watched a movie because Tony a drama.

12. (a) Kyle baked a cake, and Jack did too.
    (b) Kyle baked a cake because Jack did.
    (c) Kyle baked a cake, and Jack a cookie.
    (d) * Kyle baked a cake because Jack a cookie.

13. (a) Sara picked a pen, and Kelly can too.
    (b) Sara picked a pen because Kelly can.
    (c) Sara picked a pen, and Kelly a pencil.
    (d) * Sara picked a pen because Kelly a pencil.

14. (a) George studied math, and Sally can too.
    (b) George studied math because Sally can.
    (c) George studied math, and Sally science.
    (d) * George studied math because Sally science.

15. (a) David lost the wallet, and Kris can too.
(b) David lost the wallet because Kris can.
(c) David lost the wallet, and Kris the backpack.
(d) * David lost the wallet because Kris the backpack.

16. (a) Tom brought chocolates, and Alice can too.
(b) Tom brought chocolates because Alice can.
(c) Tom brought chocolates, and Alice candies.
(d) * Tom brought chocolates because Alice candies.

17. (a) John took forks, and Bill can too.
(b) John took forks because Bill can.
(c) John took forks, and Bill spoons.
(d) * John took forks because Bill spoons.

18. (a) Andrew saw the glasses, and Paul can too.
(b) Andrew saw the glasses because Paul can.
(c) Andrew saw the glasses, and Paul the cellphone.
(d) * Andrew saw the glasses because Paul the cellphone.

19. (a) Bob cut the carrots, and Nate can too.
(b) Bob cut the carrots because Nate can.
(c) Bob cut the carrots, and Nate the apples.
(d)* Bob cut the carrots because Nate the apples.

20. (a) Brian cleaned the desk, and Kevin can too.
(b) Brian cleaned the desk because Kevin can.
(c) Brian cleaned the desk, and Kevin the blackboard.
(d) * Brian cleaned the desk because Kevin the blackboard.

21. (a) Andy wore pants, and Helen can too.
(b) Andy wore pants because Helen can.
(c) Andy wore pants, and Helen a skirt.
(d) * Andy wore pants because Helen a skirt.

22. (a) Fred hugged the pillow, and Ryan can too.
(b) Fred hugged the pillow because Ryan can.
(c) Fred hugged the pillow, and Ryan the cushion.
(d) * Fred hugged the pillow because Ryan the cushion.

23. (a) Ann rode a bicycle, and Tony can too.
(b) Ann rode a bicycle because Tony can.
(c) Ann rode a bicycle, and Tony a motorcycle.
(d) * Ann rode a bicycle because Tony a motorcycle.

24. (a) Kyle played a card game, and Jack can too.
(b) Kyle played a card game because Jack can.
(c) Kyle played a card game, and Jack a board game.
(d) *Kyle played a card game because Jack a board game.

Filler items: null vs. overt argument

(a) No null argument, (b) *Null subject, (c) *Null object

1. (a) The ball was dirty, but I loved it.
(b) * The ball was dirty, but loved it.
(c) * The ball was dirty, but I loved.

2. (a) The laptop was heavy, but I brought it.
(b) * The laptop was heavy, but brought it.
(c) * The laptop was heavy, but I brought.

3. (a) The table was cheap, but I hated it.
(b) * The table was cheap, but hated it.
(c) * The table was cheap, but I hated.

4. (a) The book was helpful, but I sold it.
(b) * The book was helpful, but sold it.
(c) * The book was helpful, but I sold.

5. (a) The ring was shiny, so I stole it.
(b) * The ring was shiny, so stole it.
(c) * The ring was shiny, so I stole.

6. (a) The flower was beautiful, so I kept it.
(b) * The flower was beautiful, so kept it.
(c) * The flower was beautiful, so I kept.

7. (a) The paper was big, so I cut it.
(b) * The paper was big, so cut it.
(c) * The paper was big, so I cut.

8. (a) The bus was coming, so I took it.
(b) * The bus was coming, so took it.
(c) * The bus was coming, so I took.

9. (a) The movie was famous, so I watched it.
(b) * The movie was famous, so watched it.
(c) * The movie was famous, so I watched.

10. (a) The chocolate was too sweet, but you liked it.
    (b) * The chocolate was too sweet, but liked it.
    (c) * The chocolate was too sweet, but you liked.

11. (a) The bag was ugly, but you bought it.
    (b) * The bag was ugly, but bought it.
    (c) * The bag was ugly, but you bought.

12. (a) The exam was hard, but you made it.
    (b) * The exam was hard, but made it.
    (c) * The exam was hard, but you made.

13. (a) The laptop was expensive, but you wanted it.
    (b) * The laptop was expensive, but wanted it.
    (c) * The laptop was expensive, but you wanted.

14. (a) The door was heavy, but you pushed it.
    (b) * The door was heavy, but pushed it.
    (c) * The door was heavy, but you pushed.

15. (a) The ball was there, so you kicked it.
    (b) * The ball was there, so kicked it.
    (c) * The ball was there, so you kicked.

16. (a) The toy was broken, so you fixed it.
    (b) * The toy was broken, so fixed it.
    (c) * The toy was broken, so you fixed.

17. (a) The doll was very cute, so you hugged it.
    (b) * The doll was very cute, so hugged it.
    (c) * The doll was very cute, so you hugged.

18. (a) The pen was good, so you used it.
    (b) * The pen was good, so used it.
    (c) * The pen was good, so you used.

Filler items: *wanna* contraction

(a) If + No gap, (b) *If + Gap, (c) *Who + No gap, (d) Who + Gap

1. (a) I wonder if you wanna read.
   (b) * I wonder if you wanna read with.
   (c) * I wonder who you wanna read.
   (d) I wonder who you wanna read with.

2. (a) I wonder if you wanna dance.
   (b) * I wonder if you wanna dance with.
   (c) * I wonder who you wanna dance.
   (d) I wonder who you wanna dance with.

3. (a) I wonder if you wanna sing.
   (b) * I wonder if you wanna sing with.
   (c) * I wonder who you wanna sing.
   (d) I wonder who you wanna sing with.

4. (a) I wonder if you wanna walk.
   (b) * I wonder if you wanna walk with.
   (c) * I wonder who you wanna walk.
   (d) I wonder who you wanna walk with.

5. (a) I wonder if you wanna stay.
   (b) * I wonder if you wanna stay with.
   (c) * I wonder who you wanna stay.
   (d) I wonder who you wanna stay with.

6. (a) I wonder if you wanna swim.
   (b) * I wonder if you wanna swim with.
   (c) * I wonder who you wanna swim.
   (d) I wonder who you wanna swim with.

7. (a) I wonder if you wanna run.
   (b) * I wonder if you wanna run with.
   (c) * I wonder who you wanna run.
   (d) I wonder who you wanna run with.

8. (a) I wonder if you wanna smile.
   (b) * I wonder if you wanna smile with.
   (c) * I wonder who you wanna smile.
   (d) I wonder who you wanna smile with.

9. (a) I wonder if you wanna start.
   (b) * I wonder if you wanna start with.
   (c) * I wonder who you wanna start.
   (d) I wonder who you wanna start with.

10. (a) I wonder if you wanna talk.
    (b) * I wonder if you wanna talk with.
    (c) * I wonder who you wanna talk.
    (d) I wonder who you wanna talk with.

11. (a) I wonder if you wanna eat.
    (b) * I wonder if you wanna eat with.
    (c) * I wonder who you wanna eat.
    (d) I wonder who you wanna eat with.

12. (a) I wonder if you wanna travel.
    (b) * I wonder if you wanna travel with.
    (c) * I wonder who you wanna travel.
    (d) I wonder who you wanna travel with.

13. (a) I wonder if you wanna drink.
    (b) * I wonder if you wanna drink with.
    (c) * I wonder who you wanna drink.
    (d) I wonder who you wanna drink with.

14. (a) I wonder if you wanna live.
    (b) * I wonder if you wanna live with.
    (c) * I wonder who you wanna live.
    (d) I wonder who you wanna live with.

15. (a) I wonder if you wanna cry.
    (b) * I wonder if you wanna cry with.
    (c) * I wonder who you wanna cry.
    (d) I wonder who you wanna cry with.

16. (a) I wonder if you wanna laugh.
    (b) * I wonder if you wanna laugh with.
    (c) * I wonder who you wanna laugh.
    (d) I wonder who you wanna laugh with.

17. (a) I wonder if you wanna wait.
    (b) * I wonder if you wanna wait with.
    (c) * I wonder who you wanna wait.
    (d) I wonder who you wanna wait with.

18. (a) I wonder if you wanna shout.
    (b) * I wonder if you wanna shout with.
    (c) * I wonder who you wanna shout.
    (d) I wonder who you wanna shout with.

19. (a) I wonder if you wanna pray.
    (b) * I wonder if you wanna pray with.
    (c) * I wonder who you wanna pray.
    (d) I wonder who you wanna pray with.

20. (a) I wonder if you wanna work.
    (b) * I wonder if you wanna work with.
    (c) * I wonder who you wanna work.
    (d) I wonder who you wanna work with.

Filler items: missing 3sg present [–s] subject–verb agreement

1. * Tony says that his mom want a car.
2. * Ivan thinks that his wife like coffee.
3. * Sally believes that her teacher write textbooks.

Filler items: ungrammatical backward gapping pattern

1. * Jane meat, and I ate noodles.
2. * Ryan the chair, and I liked the desk.
3. * Bob the napkin, and I used the wet wipe.

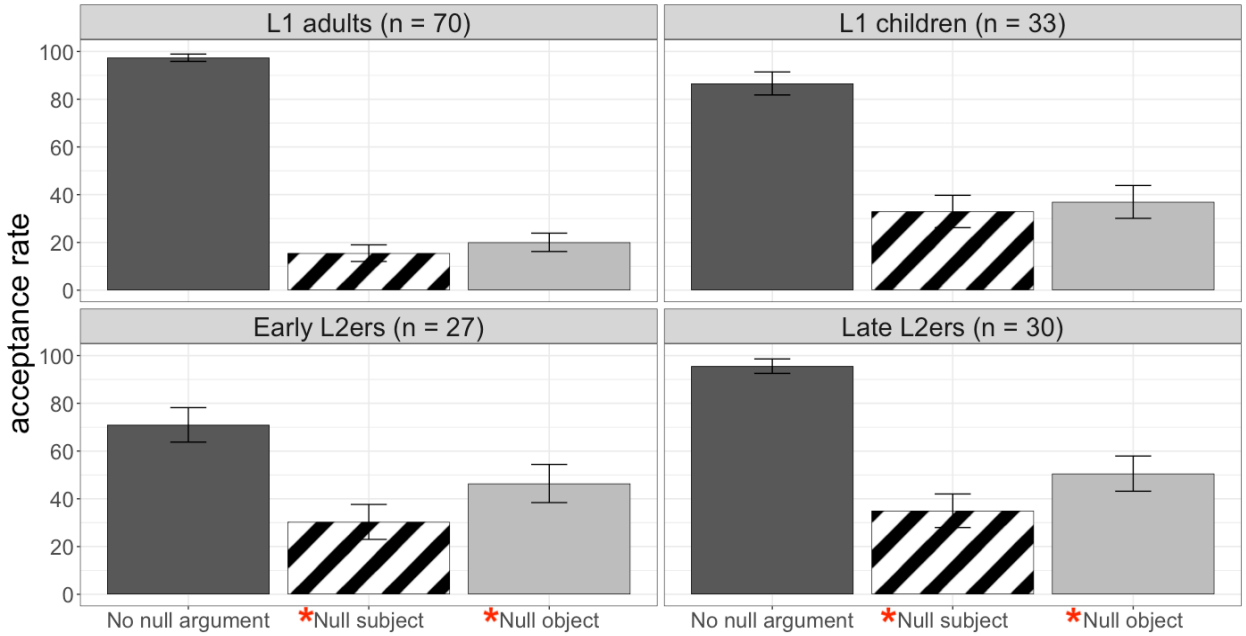
Appendix C:

Results of the Filler Items in the Acceptability Judgment Task

The mean acceptance rates for each filler type in the acceptability judgment task are presented by group in Figures C-1–C-4.

Figure C-1

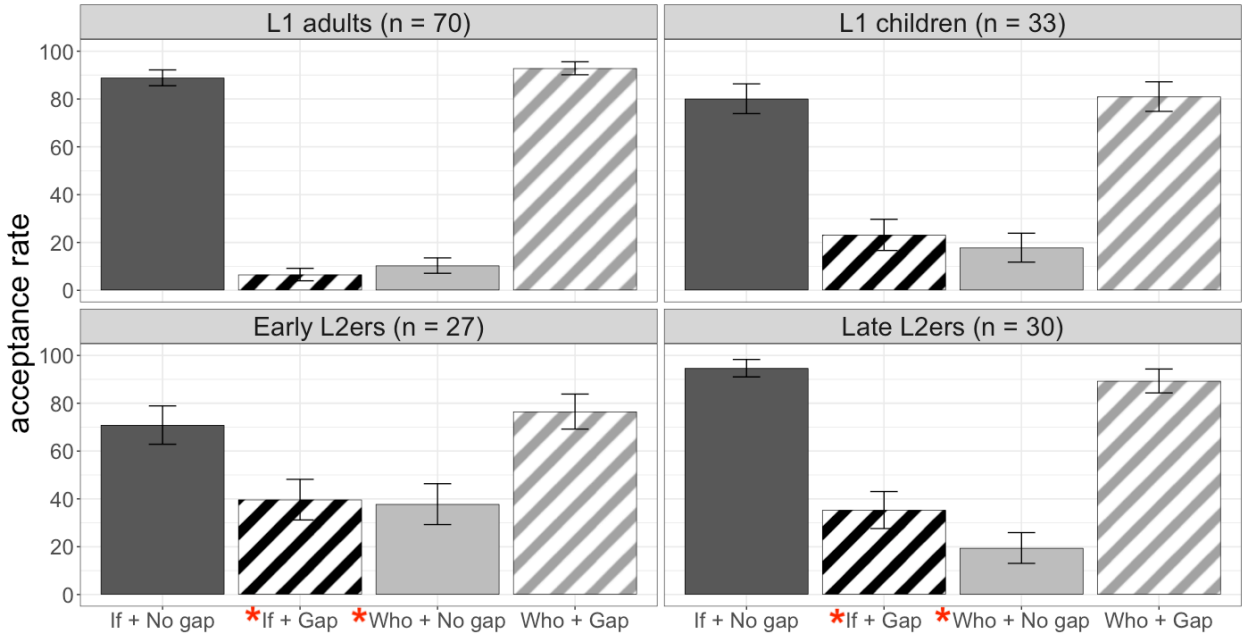
Mean Acceptance Rates (in %) of the Null vs. Overt Argument Fillers in the Acceptability Judgment Task per Condition and Group



Note. Error bars represent 95% confidence intervals.

Figure C-2

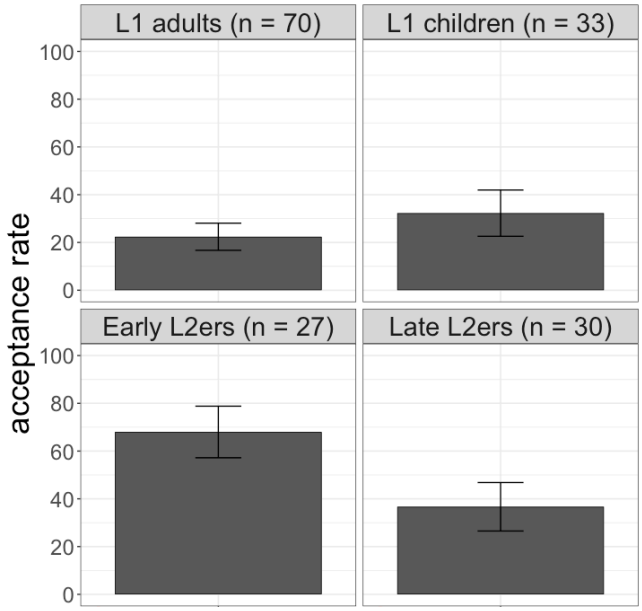
Mean Acceptance Rates (in %) of the Wanna-Contraction Fillers in the Acceptability Judgment Task per Condition and Group



Notes. If + No gap: *wanna*-contraction in the *if*-clause without a gap; *If + Gap: *wanna*-contraction in the *if*-clause with a gap; *Who + No gap: *wanna*-contraction in the *who*-clause without a gap; Who + Gap: *wanna*-contraction in the *who*-clause with a gap. Error bars represent 95% confidence intervals.

Figure C-3

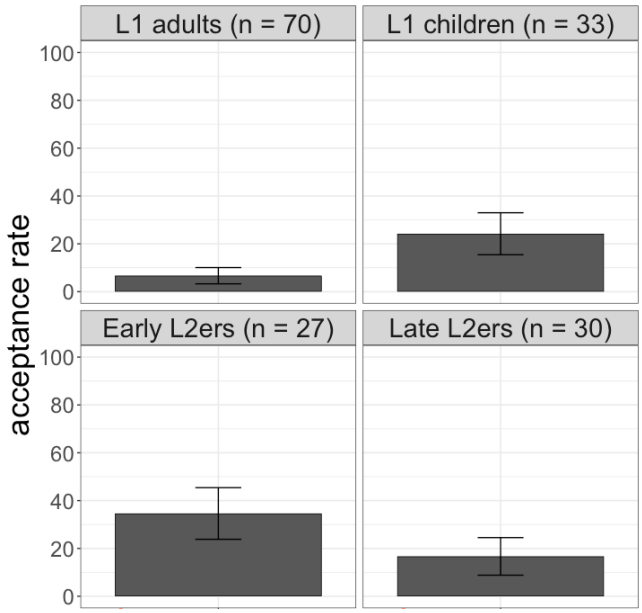
*Mean Acceptance Rates (in %) of the *Missing 3sg Present [–s] Subject–Verb Agreement Fillers in the Acceptability Judgment Task per Group*



Note. Error bars represent 95% confidence intervals.

Figure C-4

*Mean Acceptance Rates (in %) of the *Backward Gapping Fillers in the Acceptability Judgment Task per Group*



Note. Error bars represent 95% confidence intervals.

Appendix D:

Pitch Tracks for Sample Stimuli from the Acceptability Judgment Task

Each of the following figures shows the F0 contour for a sample stimulus from one of the critical conditions in the acceptability judgment task.

Figure D-1

Pitch Track for a Sample Sentence of VPE in a Conjunct Clause

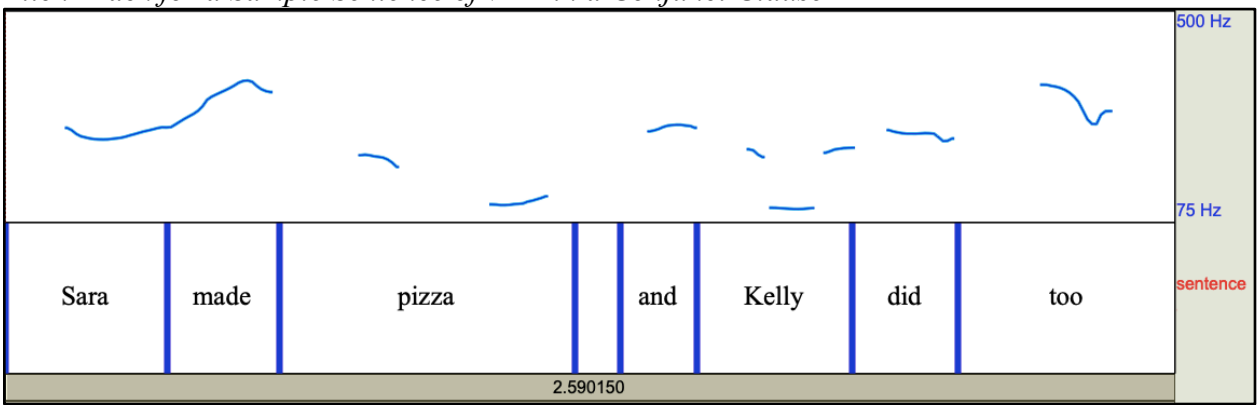


Figure D-2

Pitch Track for a Sample Sentence of VPE in an Adjunct Clause

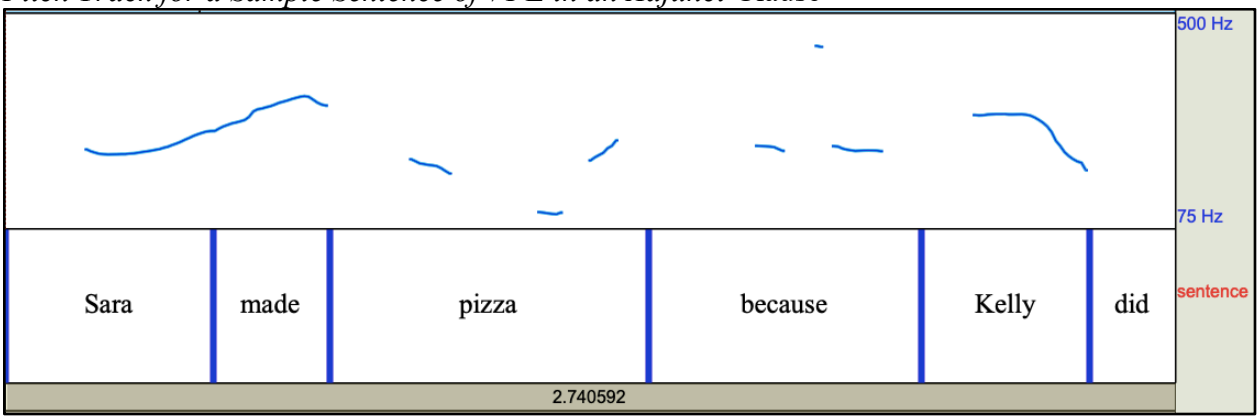


Figure D-3

Pitch Track for a Sample Sentence of Gapping in a Conjunct Clause

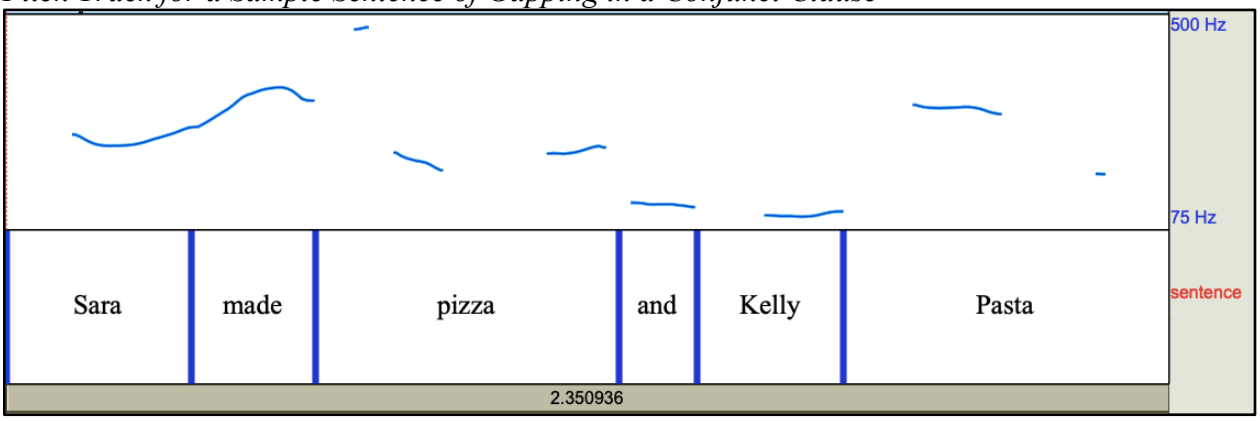
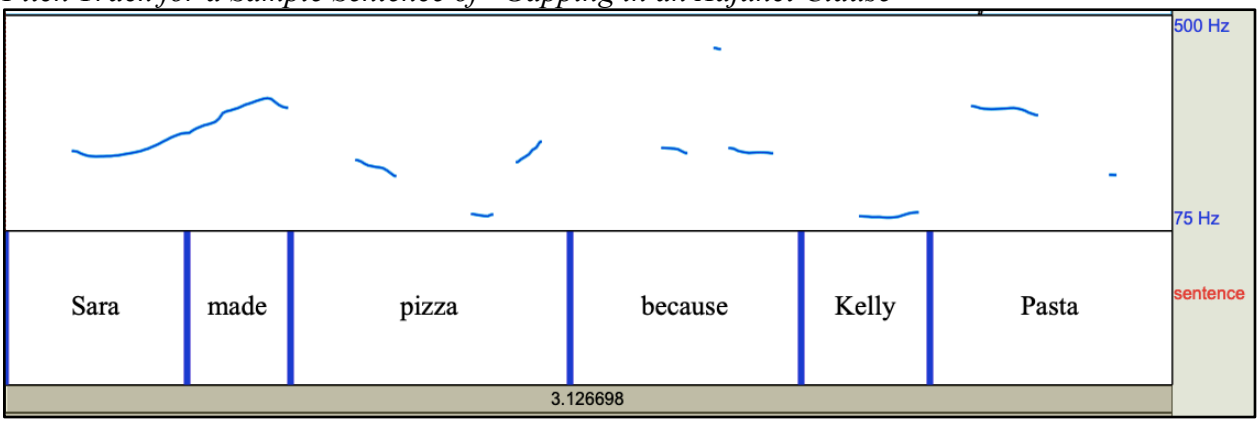


Figure D-4

*Pitch Track for a Sample Sentence of *Gapping in an Adjunct Clause*



Appendix E:

Background Questionnaires

Questionnaire for L1-English adults and late/adult L1-Korean L2 learners of English

Please answer all of the following questions.

- Participant ID:
- Age:
- Gender:
- What is your native language (the language you are firstly exposed to)?
- What is your father's native language?
- What is your mother's native language?
- At what age did you begin acquiring English?
- What language do you usually speak at home?
- If you speak more than one language at home, how much do you use English (out of 100%)?
- Please list all languages you know in order of dominance and rate your proficiency on each language from 1 to 10 in the parentheses.
- How much reading proficiency do you have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- How much writing proficiency do you have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- How much listening proficiency do you have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- How much speaking proficiency do you have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- For native English speakers only: Do you speak Hawaiian pidgin? If so, how long have you used it? (If you don't know pidgin, just put "no" in the box.)
- For second language learners of English only: If you have ever lived in or visited a country where languages other than your native language are spoken, please indicate below the name of the country, the duration and period of the stay, and which languages you used while you were in that country. (e.g., US: January 2017–June 2018, 18 months)

Questionnaire for L1-English children and early/child L1-Korean L2 learners of English
(Filled out by parents)

Please answer all of the following questions.

- Participant ID:
- Age:
- Gender:
- What is a native language of your child's mother?
- What is a native language of your child's father?
- What is the first language your child was exposed to as a child?
- What is your child's native language?
- At what age did your child begin acquiring English?
- What language does your child usually speak at home?
- If your child speaks more than one language at home, how much does he/she use English (out of 100%)?
- Please list all languages your child knows in order of dominance and rate his/her proficiency on each language from 1 to 10 in the parentheses.
- How much reading proficiency does your child have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- How much writing proficiency does your child have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- How much listening proficiency does your child have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- How much speaking proficiency does your child have with the English language? Please choose one from 1 (very low proficiency) to 10 (very high proficiency).
- For L1-English children only: Does your child speak Hawaiian pidgin? If so, how long have your child used it? (If your child doesn't know pidgin, just put "no" in the box.)
- For second language learners of English only: If your child has ever lived in or visited a country where languages other than your native language are spoken, please indicate below the name of the country, the duration and period of the stay, and which languages your child used while your child was in that country. (e.g., US: January 2017–June 2018, 18 months)

Appendix F:

Individual Sensitivity Scores in the Acceptability Judgment Task

L1-English adults (*n* = 70)

Participant code	Sensitivity score	Participant code	Sensitivity score
L1A_01	100.00	L1A_41	94.44
L1A_02	88.89	L1A_42	77.77
L1A_03	100.00	L1A_43	72.22
L1A_04	100.00	L1A_44	83.33
L1A_05	100.00	L1A_45	94.44
L1A_06	66.67	L1A_46	77.78
L1A_07	83.33	L1A_47	100.00
L1A_08	77.78	L1A_48	94.44
L1A_09	77.78	L1A_49	100.00
L1A_10	77.78	L1A_50	83.33
L1A_11	100.00	L1A_51	44.44
L1A_12	83.33	L1A_52	100.00
L1A_13	88.89	L1A_53	67.46
L1A_14	88.89	L1A_54	61.11
L1A_15	100.00	L1A_55	66.66
L1A_16	100.00	L1A_56	72.22
L1A_17	100.00	L1A_57	88.89
L1A_18	61.11	L1A_58	44.44
L1A_19	38.89	L1A_59	50.00
L1A_20	55.55	L1A_60	88.89
L1A_21	94.44	L1A_61	100.00
L1A_22	77.78	L1A_62	94.44
L1A_23	100.00	L1A_63	100.00
L1A_24	94.44	L1A_64	55.56
L1A_25	66.67	L1A_65	94.44
L1A_26	77.78	L1A_66	60.00
L1A_27	83.33	L1A_67	61.11
L1A_28	100.00	L1A_68	94.44
L1A_29	77.78	L1A_69	100.00
L1A_30	88.89	L1A_70	94.44
L1A_31	55.55		
L1A_32	83.33		
L1A_33	72.22		
L1A_34	33.33		
L1A_35	83.33		
L1A_36	87.78		
L1A_37	100.00		
L1A_38	55.56		
L1A_39	72.22		
L1A_40	94.44		

L1-English children (*n* = 33)

Participant code	Sensitivity score
L1C_01	64.44
L1C_02	100.00
L1C_03	−31.11
L1C_04	−6.67
L1C_05	16.67
L1C_06	38.88
L1C_07	61.11
L1C_08	38.89
L1C_09	81.11
L1C_10	26.67
L1C_11	50.00
L1C_12	66.67
L1C_13	41.11
L1C_14	−10.00
L1C_15	72.22
L1C_16	44.45
L1C_17	46.67
L1C_18	61.11
L1C_19	50.00
L1C_20	11.11
L1C_21	61.11
L1C_22	50.00
L1C_23	55.55
L1C_24	33.33
L1C_25	−22.22
L1C_26	66.67
L1C_27	37.78
L1C_28	−16.67
L1C_29	61.11
L1C_30	55.56
L1C_31	33.34
L1C_32	66.67
L1C_33	53.34

Early L1-Korean L2 learners of English (*n* = 27)

Participant code	Sensitivity score
EL2_01	55.56
EL2_02	44.44
EL2_03	44.44
EL2_04	20.00
EL2_05	16.67
EL2_06	67.78
EL2_07	25.00
EL2_08	−20.00
EL2_09	22.22
EL2_10	−13.33
EL2_11	19.44
EL2_12	27.78
EL2_13	50.00
EL2_14	38.89
EL2_15	55.55
EL2_16	61.11
EL2_17	72.22
EL2_18	27.78
EL2_19	53.33
EL2_20	66.67
EL2_21	55.55
EL2_22	83.33
EL2_23	27.77
EL2_24	13.33
EL2_25	55.56
EL2_26	27.78
EL2_27	32.22

Late L1-Korean L2 learners of English (*n* = 30)

Participant code	Sensitivity score
LL2_01	77.78
LL2_02	16.67
LL2_03	38.89
LL2_04	83.33
LL2_05	66.67
LL2_06	66.66
LL2_07	100.00
LL2_08	88.89
LL2_09	50.00
LL2_10	44.44
LL2_11	72.22
LL2_12	72.22
LL2_13	50.00
LL2_14	72.22
LL2_15	61.11
LL2_16	66.67
LL2_17	77.77
LL2_18	83.33
LL2_19	88.89
LL2_20	100.00
LL2_21	63.33
LL2_22	66.66
LL2_23	94.44
LL2_24	16.67
LL2_25	66.67
LL2_26	50.00
LL2_27	61.11
LL2_28	88.89
LL2_29	61.11
LL2_30	66.67

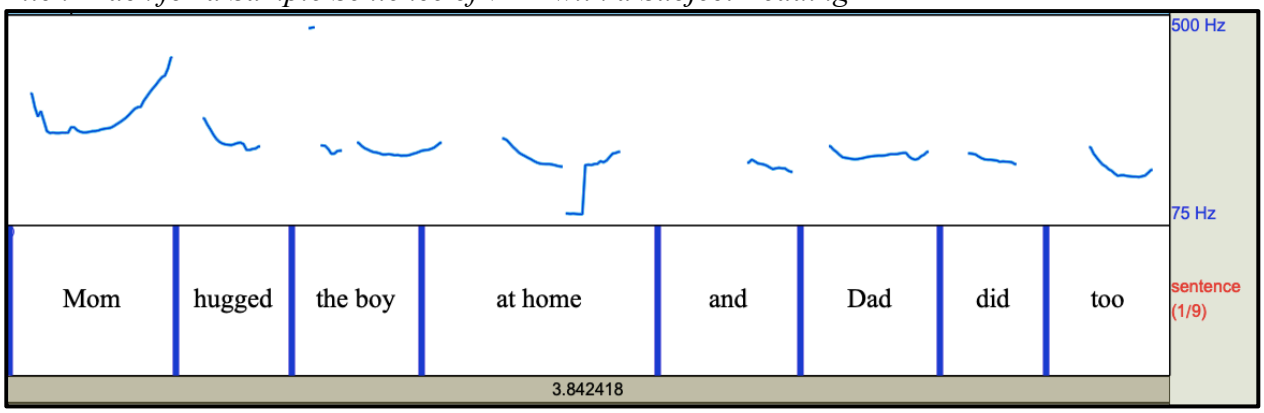
Appendix G:

Pitch Tracks for Sample Stimuli from the Picture-Sentence Matching Task

The following figures show the F0 contours for sample stimuli from each of the critical conditions in the picture-sentence matching task.

Figure G-1

Pitch Track for a Sample Sentence of VPE with a Subject Reading



Note. The same recording was used for (impossible) VPE with an object reading.

Figure G-2

*Pitch Track for a Sample Sentence of *VPE with an Object Reading*

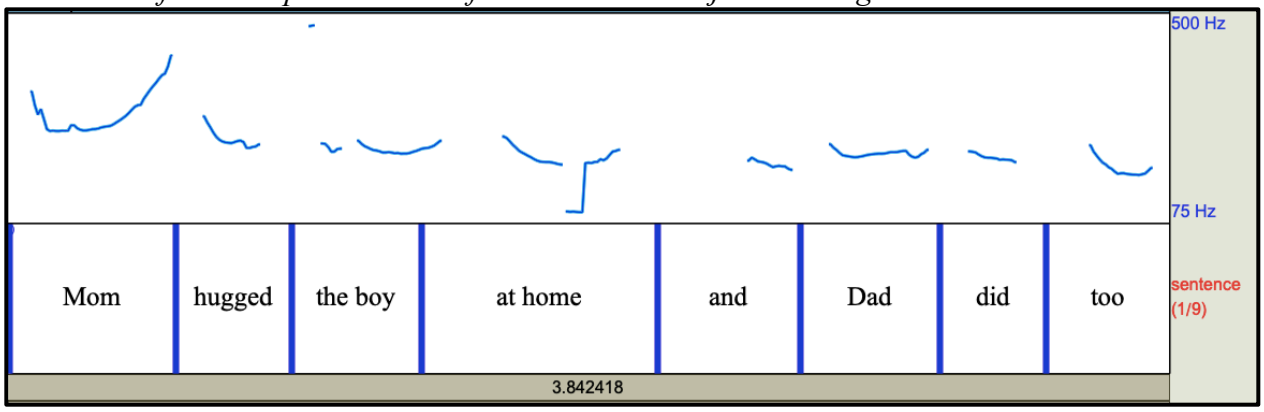


Figure G-3

Pitch Track for a Sample Sentence of Gapping with a Subject Reading

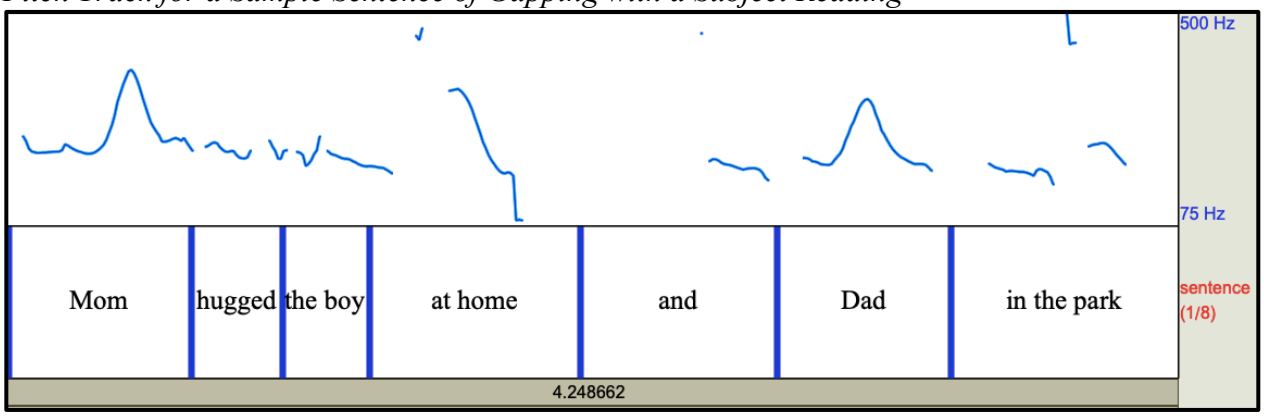
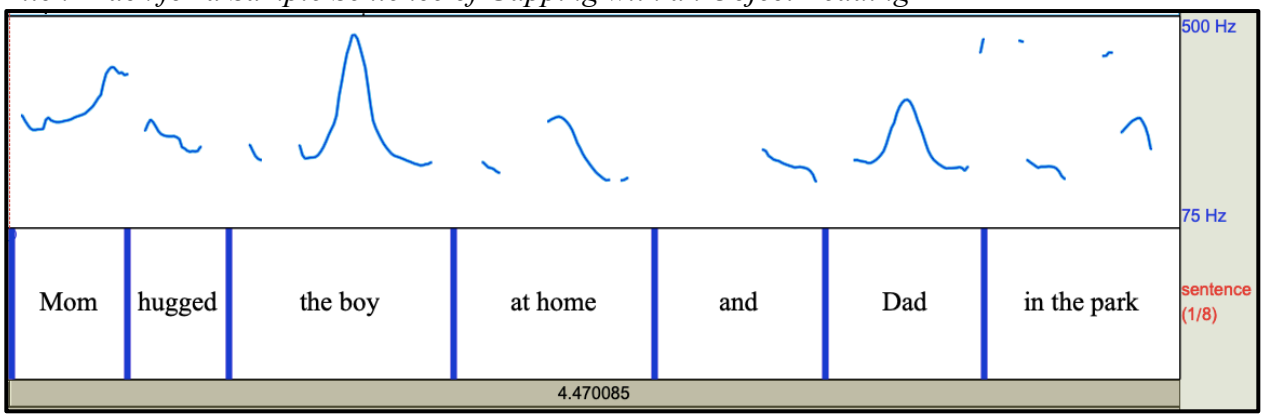


Figure G-4

Pitch Track for a Sample Sentence of Gapping with an Object Reading



Appendix H:

Stimuli for the Picture-Sentence Matching Task

Instructions:

This is a picture-sentence matching task. In this task, each trial will begin with a short video story. At the end of the story, a puppet named "Susie" will pop up on the screen and the narrator will ask her to say what happened in the story. The puppet will then say what happened. Press the smiling face on the keyboard if the puppet is right, press the frowning face if the puppet is wrong, and press the question mark if you don't know whether the puppet is right or wrong for some reason.

Critical items

1.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Mom hugged the boy at home. Dad hugged the boy at home too. Target sentence: Mom hugged the boy at home and Dad did too.
*VPE-OR (MISMATCH)		Story: Mom hugged the boy at home. Mom hugged Dad at home too. Target sentence: *Mom hugged the boy at home and Dad did too.
Gapping-SR (MATCH)		Story: Mom hugged the boy at home. Dad hugged the boy in the park. Target sentence: Mom hugged the boy at home and Dad in the park.
Gapping-OR (MATCH)		Story: Mom hugged the boy at home. Mom hugged Dad in the park. Target sentence: Mom hugged the boy at home and Dad in the park.

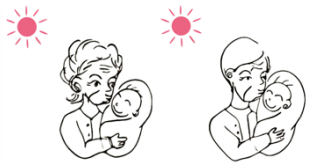
2.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Grandma hugged the kid in the morning. Grandpa hugged the kid in the morning too. Target sentence: Grandma hugged the kid in the morning and Grandpa did too.
*VPE-OR (MISMATCH)		Story: Grandma hugged the kid in the morning. Grandma hugged Grandpa in the morning too. Target sentence: *Grandma hugged the kid in the morning and Grandpa did too.
Gapping-SR (MATCH)		Story: Grandma hugged the kid in the morning. Grandpa hugged the kid in the evening. Target sentence: Grandma hugged the kid in the morning and Grandpa in the evening.
Gapping-OR (MATCH)		Story: Grandma hugged the kid in the morning. Grandma hugged Grandpa in the evening. Target sentence: Grandma hugged the kid in the morning and Grandpa in the evening.

3.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Mom kissed the little girl at home. Dad kissed the little girl at home too. Target sentence: Mom kissed the little girl at home and Dad did too.
*VPE-OR (MISMATCH)		Story: Mom kissed the little girl at home. Mom kissed Dad at home too. Target sentence: *Mom kissed the little girl at home and Dad did too.
Gapping-SR (MATCH)		Story: Mom kissed the little girl at home. Dad kissed the little girl in the park. Target sentence: Mom kissed the little girl at home and Dad in the park.
Gapping-OR (MATCH)		Story: Mom kissed the little girl at home. Mom kissed Dad in the park. Target sentence: Mom kissed the little girl at home and Dad in the park.

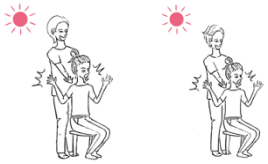
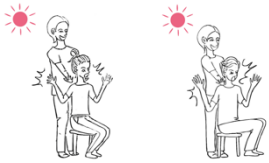
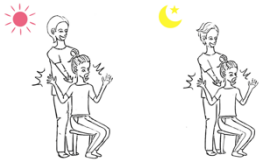
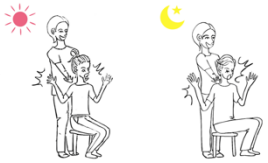
4.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Grandma kissed the baby in the morning. Grandpa kissed the baby in the morning too. Target sentence: Grandma kissed the baby in the morning and Grandpa did too.
*VPE-OR (MISMATCH)		Story: Grandma kissed the baby in the morning. Grandma kissed Grandpa in the morning too. Target sentence: *Grandma kissed the baby in the morning and Grandpa did too.
Gapping-SR (MATCH)		Story: Grandma kissed the baby in the morning. Grandpa kissed the baby in the evening. Target sentence: Grandma kissed the baby in the morning and Grandpa in the evening.
Gapping-OR (MATCH)		Story: Grandma kissed the baby in the morning. Grandma kissed Grandpa in the evening. Target sentence: Grandma kissed the baby in the morning and Grandpa in the evening.

5.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Grandma surprised the little boy at the zoo. Grandpa surprised the little boy at the zoo too. Target sentence: Grandma surprised the little boy at the zoo and Grandpa did too.
*VPE-OR (MISMATCH)		Story: Grandma surprised the little boy at the zoo. Grandma surprised Grandpa at the zoo too. Target sentence: *Grandma surprised the little boy at the zoo and Grandpa did too.
Gapping-SR (MATCH)		Story: Grandma surprised the little boy at the zoo. Grandpa surprised the little boy at the mall. Target sentence: Grandma surprised the little boy at the zoo and Grandpa at the mall.
Gapping-OR (MATCH)		Story: Grandma surprised the little boy at the zoo. Grandma surprised Grandpa at the mall. Target sentence: Grandma surprised the little boy at the zoo and Grandpa at the mall.

6.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Mom surprised the girl in the morning. Dad surprised the girl in the morning too. Target sentence: Mom surprised the girl in the morning and Dad did too.
*VPE-OR (MISMATCH)		Story: Mom surprised the girl in the morning. Mom surprised Dad in the morning too. Target sentence: *Mom surprised the girl in the morning and Dad did too.
Gapping-SR (MATCH)		Story: Mom surprised the girl in the morning. Dad surprised the girl in the evening. Target sentence: Mom surprised the girl in the morning and Dad in the evening.
Gapping-OR (MATCH)		Story: Mom surprised the girl in the morning. Mom surprised Dad in the evening. Target sentence: Mom surprised the girl in the morning and Dad in the evening.

7.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Mom visited the doctor in the morning. Dad visited the doctor in the morning too. Target sentence: Mom visited the doctor in the morning and Dad did too.
*VPE-OR (MISMATCH)		Story: Mom visited the doctor in the morning. Mom visited Dad in the morning too. Target sentence: *Mom visited the doctor in the morning and Dad did too.
Gapping-SR (MATCH)		Story: Mom visited the doctor in the morning. Dad visited the doctor in the evening. Target sentence: Mom visited the doctor in the morning and Dad in the evening.
Gapping-OR (MATCH)		Story: Mom visited the doctor in the morning. Mom visited Dad in the evening. Target sentence: Mom visited the doctor in the morning and Dad in the evening.

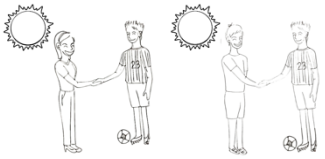
8.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Amy visited Grandma in the summer. Tom visited Grandma in the summer too. Target sentence: Amy visited Grandma in the summer and Tom did too.
*VPE-OR (MISMATCH)		Story: Amy visited Grandma in the summer. Amy visited Tom in the summer too. Target sentence: *Amy visited Grandma in the summer and Tom did too.
Gapping-SR (MATCH)		Story: Amy visited Grandma in the summer. Tom visited Grandma in the winter. Target sentence: Amy visited Grandma in the summer and Tom in the winter.
Gapping-OR (MATCH)		Story: Amy visited Grandma in the summer. Amy visited Tom in the winter. Target sentence: Amy visited Grandma in the summer and Tom in the winter.

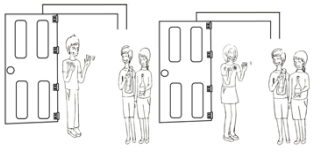
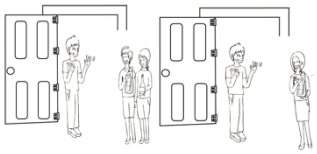
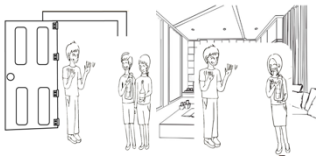
9.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Kelly met the famous singer in New York. Greg met the famous singer in New York too. Target sentence: Kelly met the famous singer in New York and Greg did too.
*VPE-OR (MISMATCH)		Story: Kelly met the famous singer in New York. Kelly met Greg in New York too. Target sentence: *Kelly met the famous singer in New York and Greg did too.
Gapping-SR (MATCH)		Story: Kelly met the famous singer in New York. Greg met the famous singer in LA. Target sentence: Kelly met the famous singer in New York and Greg in LA.
Gapping-OR (MATCH)		Story: Kelly met the famous singer in New York. Kelly met Greg in LA. Target sentence: Kelly met the famous singer in New York and Greg in LA.

10.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Tina met the soccer player in the summer. Dave met the soccer player in the summer too. Target sentence: Tina met the soccer player in the summer and Dave did too.
*VPE-OR (MISMATCH)		Story: Tina met the soccer player in the summer. Tina met Dave in the summer too. Target sentence: *Tina met the soccer player in the summer and Dave did too.
Gapping-SR (MATCH)		Story: Tina met the soccer player in the summer. Dave met the soccer player in the winter. Target sentence: Tina met the soccer player in the summer and Dave in the winter.
Gapping-OR (MATCH)		Story: Tina met the soccer player in the summer. Tina met Dave in the winter. Target sentence: Tina met the soccer player in the summer and Dave in the winter.

11.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: George greeted the guests at the front door. Alice greeted the guests at the front door too. Target sentence: George greeted the guests at the front door and Alice did too.
*VPE-OR (MISMATCH)		Story: George greeted the guests at the front door. George greeted Alice at the front door too. Target sentence: *George greeted the guests at the front door and Alice did too.
Gapping-SR (MATCH)		Story: George greeted the guests at the front door. Alice greeted the guests in the living room. Target sentence: George greeted the guests at the front door and Alice in the living room.
Gapping-OR (MATCH)		Story: George greeted the guests at the front door. George greeted Alice in the living room. Target sentence: George greeted the guests at the front door and Alice in the living room.

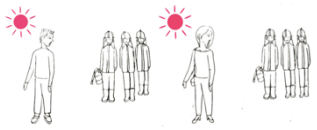
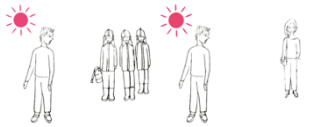
12.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Sam greeted the children on the bus. Jenny greeted the children on the bus too. Target sentence: Sam greeted the children on the bus and Jenny did too.
*VPE-OR (MISMATCH)		Story: Sam greeted the children on the bus. Sam greeted Jenny on the bus too. Target sentence: *Sam greeted the children on the bus and Jenny did too.
Gapping-SR (MATCH)		Story: Sam greeted the children on the bus. Jenny greeted the children at the zoo. Target sentence: Sam greeted the children on the bus and Jenny at the zoo.
Gapping-OR (MATCH)		Story: Sam greeted the children on the bus. Sam greeted Jenny at the zoo. Target sentence: Sam greeted the children on the bus and Jenny at the zoo.

13.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Jane saw the students at the zoo. Fred saw the students at the zoo too. Target sentence: Jane saw the students at the zoo and Fred did too.
*VPE-OR (MISMATCH)		Story: Jane saw the students at the zoo. Jane saw Fred at the zoo too. Target sentence: *Jane saw the students at the zoo and Fred did too.
Gapping-SR (MATCH)		Story: Jane saw the students at the zoo. Fred saw the students in the park. Target sentence: Jane saw the students at the zoo and Fred in the park.
Gapping-OR (MATCH)		Story: Jane saw the students at the zoo. Jane saw Fred in the park. Target sentence: Jane saw the students at the zoo and Fred in the park.

14.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Brad saw some firefighters in the morning. Emma saw some firefighters in the morning too. Target sentence: Brad saw some firefighters in the morning and Emma did too.
*VPE-OR (MISMATCH)		Story: Brad saw some firefighters in the morning. Brad saw Emma in the morning too. Target sentence: *Brad saw some firefighters in the morning and Emma did too.
Gapping-SR (MATCH)		Story: Brad saw some firefighters in the morning. Emma saw some firefighters in the evening. Target sentence: Brad saw some firefighters in the morning and Emma in the evening.
Gapping-OR (MATCH)		Story: Brad saw some firefighters in the morning. Brad saw Emma in the evening. Target sentence: Brad saw some firefighters in the morning and Emma in the evening.

15.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Mom called the doctor from home. Dad called the doctor from home too. Target sentence: Mom called the doctor from home and Dad did too.
*VPE-OR (MISMATCH)		Story: Mom called the doctor from home. Mom called Dad from home too. Target sentence: *Mom called the doctor from home and Dad did too.
Gapping-SR (MATCH)		Story: Mom called the doctor from home. Dad called the doctor from the office. Target sentence: Mom called the doctor from home and Dad from the office.
Gapping-OR (MATCH)		Story: Mom called the doctor from home. Mom called Dad from the office. Target sentence: Mom called the doctor from home and Dad from the office.

16.

Condition	Pictures	Audio stimuli
VPE-SR (MATCH)		Story: Sarah called the police in the morning. Bob called the police in the morning too. Target sentence: Sarah called the police in the morning and Bob did too.
*VPE-OR (MISMATCH)		Story: Sarah called the police in the morning. Sarah called Bob in the morning too. Target sentence: *Sarah called the police in the morning and Bob did too.
Gapping-SR (MATCH)		Story: Sarah called the police in the morning. Bob called the police in the evening. Target sentence: Sarah called the police in the morning and Bob in the evening.
Gapping-OR (MATCH)		Story: Sarah called the police in the morning. Sarah called Bob in the evening. Target sentence: Sarah called the police in the morning and Bob in the evening.

Filler items: *verb mismatch

1.

Condition	Pictures	Audio stimuli
*Verb mismatch (MISMATCH)		Story: The woman liked yellow flowers. The man hated red flowers. Target sentence: The woman liked yellow flowers and the man red flowers.

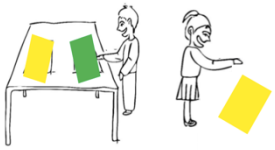
2.

Condition	Pictures	Audio stimuli
*Verb mismatch (MISMATCH)		Story: The man hated red caps. The woman liked blue caps. Target sentence: The man hated red caps, and the woman blue caps.

3.

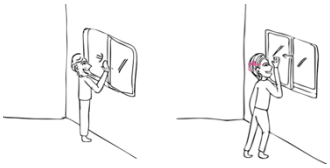
Condition	Pictures	Audio stimuli
*Verb mismatch (MISMATCH)		Story: Christina loved red dresses. Nick hated white dresses. Target sentence: Christina loved red dresses, and Nick white dresses.

4.

Condition	Pictures	Audio stimuli
*Verb mismatch (MISMATCH)		Story: George picked green paper. Alice dropped yellow paper. Target sentence: George picked green paper, and Alice yellow paper.

Filler items: *object mismatch

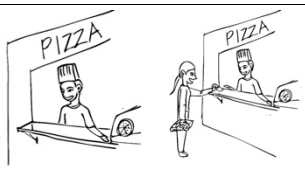
1.

Condition	Pictures	Audio stimuli
*Object mismatch (MISMATCH)		Story: Kyle opened the window. Helen closed the window. Target sentence: Kyle opened the window and Helen closed the door.

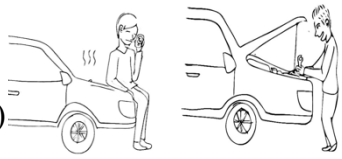
2.

Condition	Pictures	Audio stimuli
*Object mismatch (MISMATCH)		Story: Robert bought the car. Lihana loved the car. Target sentence: Robert bought the car, and Lihana loved the bike.

3.

Condition	Pictures	Audio stimuli
*Object mismatch (MISMATCH)		Story: Peter sold pizza. Nancy bought pizza. Target sentence: Peter sold pizza, and Nancy bought pasta.

4.

Condition	Pictures	Audio stimuli
*Object mismatch (MISMATCH)		Story: Jason broke the car. Fred fixed the car. Target sentence: Jason broke the car, and Fred fixed the computer.

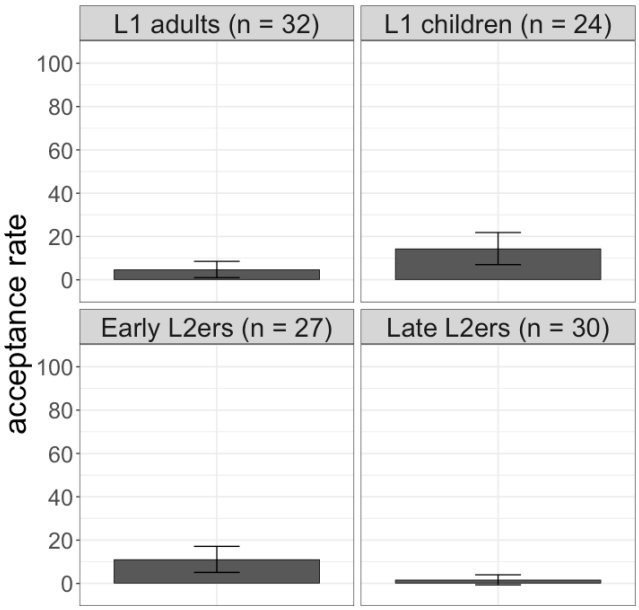
Appendix I:

Results of the Filler Items in the Picture-Sentence Matching Task

The mean acceptance rates for each filler type in the picture-sentence matching task are presented in Figures I-1 and I-2.

Figure I-1

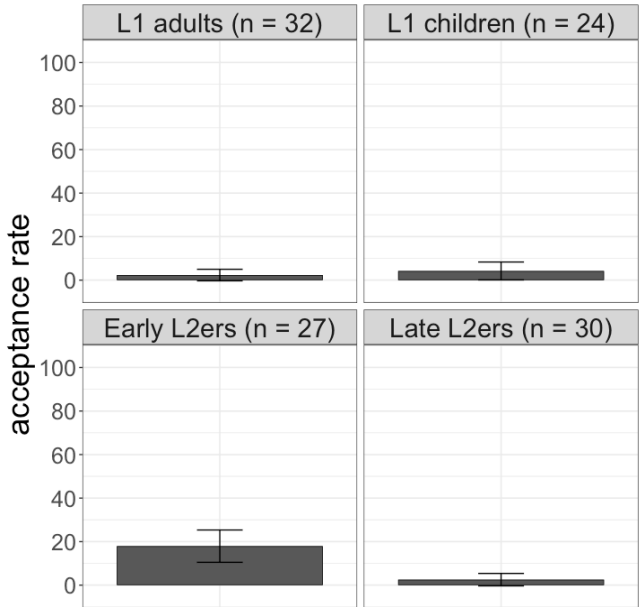
*Mean Acceptance Rates (in %) of the *Verb Mismatch Fillers in the Picture-Sentence Matching Task per Condition and Group*



Note. Error bars represent 95% confidence intervals.

Figure I-2

*Mean Acceptance Rates (in %) of the *Object Mismatch Fillers in the Picture-Sentence Matching Task per Condition and Group*



Note. Error bars represent 95% confidence intervals.

Appendix J:

Individual Sensitivity Scores in the Picture-Sentence Matching Task

L1-English adults (*n* = 32)

Participant code	Sensitivity score
L1A_39	75.00
L1A_40	91.67
L1A_41	83.33
L1A_42	100.00
L1A_43	66.67
L1A_44	75.00
L1A_45	100.00
L1A_46	75.00
L1A_47	66.67
L1A_48	91.67
L1A_49	75.00
L1A_50	100.00
L1A_51	91.67
L1A_52	91.67
L1A_53	100.00
L1A_54	50.00
L1A_55	100.00
L1A_56	58.33
L1A_57	75.00
L1A_58	75.00
L1A_59	83.33
L1A_60	100.00
L1A_61	100.00
L1A_62	83.33
L1A_63	83.33
L1A_64	66.67
L1A_65	83.33
L1A_66	77.78
L1A_67	66.67
L1A_68	66.67
L1A_69	91.67
L1A_70	58.33

L1-English children (*n* = 24)

Participant code	Sensitivity score
L1C_10	8.33
L1C_11	0.00
L1C_12	100.00
L1C_13	25.00
L1C_14	63.89
L1C_15	100.00
L1C_16	83.33
L1C_17	33.33
L1C_18	−8.33
L1C_19	41.67
L1C_20	50.00
L1C_21	66.67
L1C_22	66.67
L1C_23	75.00
L1C_24	75.00
L1C_25	16.67
L1C_26	50.00
L1C_27	91.67
L1C_28	91.67
L1C_29	83.33
L1C_30	91.67
L1C_31	41.67
L1C_32	50.00
L1C_33	75.00

Early L1-Korean L2 learners of English (*n* = 27)

Participant code	Sensitivity score
EL2_01	33.33
EL2_02	50.00
EL2_03	0.00
EL2_04	0.00
EL2_05	25.00
EL2_06	66.67
EL2_07	16.67
EL2_08	25.00
EL2_09	100.00
EL2_10	0.00
EL2_11	−25.00
EL2_12	8.33
EL2_13	83.33
EL2_14	83.33
EL2_15	100.00
EL2_16	91.67
EL2_17	91.67
EL2_18	0.00
EL2_19	50.00
EL2_20	66.67
EL2_21	66.67
EL2_22	100.00
EL2_23	16.67
EL2_24	−25.00
EL2_25	33.33
EL2_26	66.67
EL2_27	58.33

Late L1-Korean L2 learners of English (*n* = 30)

Participant code	Sensitivity score
LL2_01	33.33
LL2_02	100.00
LL2_03	100.00
LL2_04	50.00
LL2_05	91.67
LL2_06	100.00
LL2_07	75.00
LL2_08	100.00
LL2_09	91.67
LL2_10	75.00
LL2_11	91.67
LL2_12	77.78
LL2_13	66.67
LL2_14	100.00
LL2_15	75.00
LL2_16	83.33
LL2_17	83.33
LL2_18	100.00
LL2_19	100.00
LL2_20	75.00
LL2_21	75.00
LL2_22	50.00
LL2_23	83.33
LL2_24	91.67
LL2_25	75.00
LL2_26	100.00
LL2_27	100.00
LL2_28	66.67
LL2_29	75.00
LL2_30	58.33

Appendix K:
J. D. Brown's (1980) Cloze Test

Instructions:

1. Read the passage quickly to get the general meaning.
2. Write only one word in each blank. Contractions (example: don't) and possessives (John's bicycle) are one word.

Note: Spelling will not count against you as long as the scorer can read the word.

Example:
The boy walked up the street. He stepped on a piece of ice.
He fell _*down*_ but he didn't hurt himself.

MAN AND HIS PROGRESS

Man is the only living creature that can make and use tools. He is the most teachable of living beings, earning the name of Homo sapiens. (1)____ ever restless brain has used the (2)_____ and the wisdom of his ancestors (3)_____ improve his way of life. Since (4)____ is able to walk and run (5)_____ his feet, his hands have always (6)_____ free to carry and to use (7)_____. Man's hands have served him well (8)_____ his life on earth. His development, (9)_____ can be divided into three major (10)_____, is marked by several different ways (11)_____ life.

Up to 10,000 years ago, (12)_____ human beings lived by hunting and (13)_____. They also picked berries and fruits, (14)_____ dug for various edible roots. Most (15)_____, the men were the hunters, and (16)_____ women acted as food gatherers. Since (17)_____ women were busy with the children, (18)_____ men handled the tools. In a (19)_____ hand, a dead branch became a (20)_____ to knock down fruit or to (21)_____ for tasty roots. Sometimes, an animal (22)_____ served as a club, and a (23)_____ piece of stone, fitting comfortably into (24)_____ hand, could be used to break (25)_____ or to throw at an animal. (26)_____ stone was

chipped against another until (27)_____ had a sharp edge. The primitive (28)_____ who first thought of putting a (29)_____ stone at the end of a (30)_____ made a brilliant discovery: he (31)_____ joined two things to make a (32)_____ useful tool, the spear. Flint, found (33)_____ many rocks, became a common cutting (34)_____ in the Paleolithic period of man's (35)_____. Since no wood or bone tools (36)_____ survived, we know of this man (37)_____ his stone implements, with which he (38)_____ kill animals, cut up the meat, (39)_____ scrape the skins, as well as (40)_____ pictures on the walls of the (41)_____ where he lived during the winter.

(42)_____ the warmer seasons, man wandered on (43)_____ steppes of Europe without a fixed (44)_____, always foraging for food. Perhaps the (45)_____ carried nuts and berries in shells (46)_____ skins or even in light, woven (47)_____. Whenever they camped, the primitive people (48)_____ fires by striking flint for sparks (49)_____ using dried seeds, moss, and rotten (50)_____ for tinder. With fires that he kindled himself, man could keep wild animals away and could cook those that he killed, as well as provide warmth and light for himself.

Appendix L:

Stimuli for the Self-Paced Reading Task

Instructions:

This is a sentence reading task. In this task, you will see a series of dashes representing the words within a sentence. When you press the spacebar, you will see the next word in the sentence. After each sentence, you will be asked a simple question about the sentence. Your task is to read the sentence as quickly as possible, while understanding it, as checked by the question. The computer will record the time taken for each press of the spacebar.

Please click the button below to begin the task.

Critical items

(a) Gapping-Plausible, (b) *Gapping-Implausible, (c) VPE-Plausible (control),
(d) VPE-Implausible (control)

Note: The VPE-Implausible condition sentences are not themselves implausible, but use the same verbs as the Gapping-Implausible condition sentences as a control.

1. (a) Bill ordered coffee and tea at the cafe, and Jane sandwiches and cake at the bakery.
   (b) *Bill drank coffee and tea at the cafe, and Jane sandwiches and cake at the bakery.
   (c) Bill ordered coffee and tea at the cafe, and Jane did too with her brother.
   (d) Bill drank coffee and tea at the cafe, and Jane did too with her brother.

2. (a) John had coffee and tea at the cafe, and Rachel bagels and muffins at the store.
   (b) *John spilled coffee and tea at the cafe, and Rachel bagels and muffins at the store.
   (c) John had coffee and tea at the cafe, and Rachel did too with her date.
   (d) John spilled coffee and tea at the cafe, and Rachel did too with her date.

3. (a) Joe found cups and glasses at the market, and Sue pillows and blankets at the mall.
   (b) *Joe broke cups and glasses at the market, and Sue pillows and blankets at the mall.
   (c) Joe found cups and glasses at the market, and Sue did too with her neighbor.
   (d) Joe broke cups and glasses at the market, and Sue did too with her neighbor.

4. (a) Bonnie bought clothes and shoes in Paris, and Sally chocolates and cookies in Tokyo.
   (b) *Bonnie designed clothes and shoes in Paris, and Sally chocolates and cookies in Tokyo.
   (c) Bonnie bought clothes and shoes in Paris, and Sally did too with her sister.
   (d) Bonnie designed clothes and shoes in Paris, and Sally did too with her sister.

5. (a) William heated brownies and cake at home, and Troy milk and coffee at the office.
   (b) * William baked brownies and cake at home, and Troy milk and coffee at the office.
   (c) William heated brownies and cake at home, and Troy did too with his aunt.
   (d) William baked brownies and cake at home, and Troy did too with his aunt.

6. (a) Peter served steak and shrimp at the restaurant, and Liz beer and wine at the bar.
   (b) * Peter cooked steak and shrimp at the restaurant, and Liz beer and wine at the bar.
   (c) Peter served steak and shrimp at the restaurant, and Liz did too with her supervisor.
   (d) Peter cooked steak and shrimp at the restaurant, and Liz did too with her supervisor.

7. (a) Jeff cleaned the doors and the windows in the living room, and Amy the lamps and the chairs in the bedroom.
   (b) * Jeff closed the doors and the windows in the living room, and Amy the lamps and the chairs in the bedroom.
   (c) Jeff cleaned the doors and the windows in the living room, and Amy did too with her son.
   (d) Jeff closed the doors and the windows in the living room, and Amy did too with her son.

8. (a) Bob replaced the doors and the windows in the bedroom, and Nancy the fans and the chairs in the kitchen.
   (b) * Bob opened the doors and the windows in the bedroom, and Nancy the fans and the chairs in the kitchen.
   (c) Bob replaced the doors and the windows in the bedroom, and Nancy did too with her husband.
   (d) Bob opened the doors and the windows in the bedroom, and Nancy did too with her husband.

9. (a) Paul painted the cabinet and the drawer in the bedroom, and Kelly the door and the wall in the bathroom.
   (b) * Paul emptied the cabinet and the drawer in the bedroom, and Kelly the door and the wall in the bathroom.
   (c) Paul painted the cabinet and the drawer in the bedroom, and Kelly did too with her niece.
   (d) Paul emptied the cabinet and the drawer in the bedroom, and Kelly did too with her niece.

10. (a) Kevin grabbed the pill and the water at the hospital, and Mary the money and the card at the bank.
    (b) * Kevin swallowed the pill and the water at the hospital, and Mary the money and the card at the bank.
    (c) Kevin grabbed the pill and the water at the hospital, and Mary did too with her mother.
    (d) Kevin swallowed the pill and the water at the hospital, and Mary did too with her mother.

11. (a) Alice discovered cups and plates in the kitchen, and Ted stamps and cards in the bedroom.
    (b) * Alice washed cups and plates in the kitchen, and Ted stamps and cards in the bedroom.
    (c) Alice discovered cups and plates in the kitchen, and Ted did too with his cousin.
    (d) Alice washed cups and plates in the kitchen, and Ted did too with his cousin.

12. (a) Katie observed the apples and the bananas in the garden, and Brad the moon and the stars in the yard.
    (b) * Katie ate the apples and the bananas in the garden, and Brad the moon and the stars in the yard.
    (c) Katie observed the apples and the bananas in the garden, and Brad did too with his wife.
    (d) Katie ate the apples and the bananas in the garden, and Brad did too with his wife.

13. (a) Kay prepared the apples and the onions in the kitchen, and George the spoons and the forks in the dining room.
    (b) * Kay peeled the apples and the onions in the kitchen, and George the spoons and the forks in the dining room.
    (c) Kay prepared the apples and the onions in the kitchen, and George did too with the chef.
    (d) Kay peeled the apples and the onions in the kitchen, and George did too with the chef.

14. (a) Linda drew the spider and the ant in the park, and Tom the car and the truck in the classroom.
    (b) * Linda killed the spider and the ant in the park, and Tom the car and the truck in the classroom.
    (c) Linda drew the spider and the ant in the park, and Tom did too with his friend.
    (d) Linda killed the spider and the ant in the park, and Tom did too with his friend.

15. (a) Christina dropped the carrots and the potatoes in the kitchen, and James the bowls and the plates in the dining room.
    (b) * Christina cut the carrots and the potatoes in the kitchen, and James the bowls and the plates in the dining room.
    (c) Christina dropped the carrots and the potatoes in the kitchen, and James did too with the waitress.
    (d) Christina cut the carrots and the potatoes in the kitchen, and James did too with the waitress.

16. (a) Emily received the receipt and the business card at the mall, and Fred the candy and the chocolate at the restaurant.
    (b) * Emily ripped the receipt and the business card at the mall, and Fred the candy and the chocolate at the restaurant.
    (c) Emily received the receipt and the business card at the mall, and Fred did too with his colleague.
    (d) Emily ripped the receipt and the business card at the mall, and Fred did too with his colleague.

17. (a) Anna owned a car and a truck in Texas, and Kris a house and a restaurant in Boston.
    (b) *Anna drove a car and a truck in Texas, and Kris a house and a restaurant in Boston.
    (c) Anna owned a car and a truck in Texas, and Kris did too with his family.
    (d) Anna drove a car and a truck in Texas, and Kris did too with his family.

18. (a) Barbara admired the trees and the flowers in the garden, and Eric the dogs and the cats in the yard.
    (b) *Barbara watered the trees and the flowers in the garden, and Eric the dogs and the cats in the yard.
    (c) Barbara admired the trees and the flowers in the garden, and Eric did too with his daughter.
    (d) Barbara watered the trees and the flowers in the garden, and Eric did too with his daughter.

19. (a) Helen sold the piano and the flute at the mall, and Thomas the bike and the helmet at the market.
    (b) *Helen played the piano and the flute at the mall, and Thomas the bike and the helmet at the market.
    (c) Helen sold the piano and the flute at the mall, and Thomas did too with his coworker.
    (d) Helen played the piano and the flute at the mall, and Thomas did too with his coworker.

20. (a) Julie enjoyed the movie and the TV drama at home, and Tony the novel and the magazine at the library.
    (b) *Julie watched the movie and the TV drama at home, and Tony the novel and the magazine at the library.
    (c) Julie enjoyed the movie and the TV drama at home, and Tony did too with his roommate.
    (d) Julie watched the movie and the TV drama at home, and Tony did too with his roommate.

Filler items: Right Node Raising

(a) Right Node Raising-Plausible, (b) *Right Node Raising-Implausible,
(c) Right Node Raising-Plausible (control), (d) Right Node Raising-Implausible (control)

Note: The Right Node Raising-Implausible (control) condition sentences are not themselves implausible, but use the same verbs as the Right Node Raising-Implausible condition sentences as a control.

1. (a) Bill made and John sold chairs and tables during the vacation, according to their mother.
    (b) *Bill cooked and John sold chairs and tables during the vacation, according to their mother.
    (c) Bill made and John sold pizza and pasta during the vacation, according to their mother.
    (d) Bill cooked and John sold pizza and pasta during the vacation, according to their mother.

2. (a) Sam found and Andy used stamps and cards this morning, according to their mother.
   (b) * Sam washed and Andy used stamps and cards this morning, according to their mother.
   (c) Sam found and Andy used plates and bowls this morning, according to their mother.
   (d) Sam washed and Andy used plates and bowls this morning, according to their mother.

3. (a) John reviewed and Peter improved the tools and the machines three weeks ago, according to their mother.
   (b) * John wrote and Peter improved the tools and the machines three weeks ago, according to their mother.
   (c) John reviewed and Peter improved the novels and the essays three weeks ago, according to their mother.
   (d) John wrote and Peter improved the novels and the essays three weeks ago, according to their mother.

4. (a) Tom opened and Jimmy cleaned the doors and windows this evening, according to their mother.
   (b) * Tom emptied and Jimmy cleaned the doors and windows this evening, according to their mother.
   (c) Tom opened and Jimmy cleaned the drawers and cabinets this evening, according to their mother.
   (d) Tom emptied and Jimmy cleaned the drawers and cabinets this evening, according to their mother.

5. (a) Peter froze and Brad defrosted the milk and the water last night, according to their mother.
   (b) * Peter seasoned and Brad defrosted the milk and the water last night, according to their mother.
   (c) Peter froze and Brad defrosted the beef and the pork last night, according to their mother.
   (d) Peter seasoned and Brad defrosted the beef and the pork last night, according to their mother.

6. (a) Nicole studied and Claire taught math and science all day long, according to their father.
   (b) * Nicole spoke and Claire taught math and science all day long, according to their father.
   (c) Nicole studied and Claire taught English and French all day long, according to their father.
   (d) Nicole spoke and Claire taught English and French all day long, according to their father.

7. (a) Bruno painted and Mark drew butterflies and dragonflies on Sunday, according to their father.
   (b) * Bruno pruned and Mark drew butterflies and dragonflies on Sunday, according to their father.
   (c) Bruno painted and Mark drew roses and tulips on Sunday, according to their father.
   (d) Bruno pruned and Mark drew roses and tulips on Sunday, according to their father.

8. (a) Thomas brought and James ate soup and pasta this evening, according to their father.
   (b) * Thomas sliced and James ate soup and pasta this evening, according to their father.
   (c) Thomas brought and James ate apples and carrots this evening, according to their father.
   (d) Thomas sliced and James ate apples and carrots this evening, according to their father.

9. (a) Kris taught and George played football and baseball for a while, according to their father.
   (b) * Kris tuned and George played football and baseball for a while, according to their father.
   (c) Kris taught and George played the guitar and the violin for a while, according to their father.
   (d) Kris tuned and George played the guitar and the violin for a while, according to their father.

10. (a) Ben purchased and William rode a horse and a donkey an hour ago, according to their father.
    (b) * Ben parked and William rode a horse and a donkey an hour ago, according to their father.
    (c) Ben purchased and William rode a car and a bike an hour ago, according to their father.
    (d) Ben parked and William rode a car and a bike an hour ago, according to their father.

11. (a) Sue advertised and Kerry ordered the computers and mouses two weeks ago, according to their sister.
    (b) * Sue finished and Kerry ordered the computers and mouses two weeks ago, according to their sister.
    (c) Sue advertised and Kerry ordered the magazines and comic books two weeks ago, according to their sister.
    (d) Sue finished and Kerry ordered the magazines and comic books two weeks ago, according to their sister.

12. (a) Irene missed and Mary caught the train and bus last night, according to their sister.
    (b) * Irene fed and Mary caught the train and bus last night, according to their sister.
    (c) Irene missed and Mary caught the dog and cat last night, according to their sister.
    (d) Irene fed and Mary caught the dog and cat last night, according to their sister.

13. (a) Helen rated and Amy reviewed movies and TV dramas during the summer, according to their sister.
    (b) * Helen designed and Amy reviewed movies and TV dramas during the summer, according to their sister.
    (c) Helen rated and Amy reviewed computers and laptops during the summer, according to their sister.
    (d) Helen designed and Amy reviewed computers and laptops during the summer, according to their sister.

14. (a) Jane bought and Nina sold the cars and bikes during the winter, according to their sister.
    (b) * Jane watered and Nina sold the cars and bikes during the winter, according to their sister.
    (c) Jane bought and Nina sold the plants and flowers during the winter, according to their sister.
    (d) Jane watered and Nina sold the plants and flowers during the winter, according to their sister.

15. (a) Chelsea took and Nancy ate sandwiches and cookies this morning, according to their sister.
    (b) * Chelsea boiled and Nancy ate sandwiches and cookies this morning, according to their sister.
    (c) Chelsea took and Nancy ate soup and potatoes this morning, according to their sister.
    (d) Chelsea boiled and Nancy ate soup and potatoes this morning, according to their sister.

16. (a) Anna handwrote and Mariah typed the article and essay an hour ago, according to their brother.
    (b) * Anna yelled and Mariah typed the article and essay an hour ago, according to their brother.
    (c) Anna handwrote and Mariah typed the name and address an hour ago, according to their brother.
    (d) Anna yelled and Mariah typed the name and address an hour ago, according to their brother.

17. (a) Christina ordered and Mia installed the air conditioner and heater last month, according to their brother.
    (b) * Christina downloaded and Mia installed the air conditioner and heater last month, according to their brother.
    (c) Christina ordered and Mia installed the software and data last month, according to their brother.
    (d) Christina downloaded and Mia installed the software and data last month, according to their brother.

18. (a) Audrey wanted and Holly bought chocolates and cookies a few hours ago, according to their brother.
    (b) * Audrey drank and Holly bought chocolates and cookies a few hours ago, according to their brother.
    (c) Audrey wanted and Holly bought beer and wine a few hours ago, according to their brother.
    (d) Audrey drank and Holly bought beer and wine a few hours ago, according to their brother.

19. (a) Alice received and Crystal packed the drinks and snacks this morning, according to their brother.
    (b) * Alice folded and Crystal packed the drinks and snacks this morning, according to their brother.
    (c) Alice received and Crystal packed the towels and clothes this morning, according to their brother.
    (d) Alice folded and Crystal packed the towels and clothes this morning, according to their brother.

20. (a) Harry damaged and Kenny repaired the school and the library this month, according to their brother.
    (b) * Harry broke and Kenny repaired the school and the library this month, according to their brother.
    (c) Harry damaged and Kenny repaired the window and the door this month, according to their brother.
    (d) Harry broke and Kenny repaired the window and the door this month, according to their brother.

Filler items: subject–verb number agreement

(a) Article-Plural, (b) *Article- Singular, (c) Demonstrative-Plural, (d) *Demonstrative-Singular

Note. These fillers are slightly modified versions of the sentences used by Jiang (2004).

1. (a) The fires in the apartment were caused by a cigarette butt thrown on the carpet.
   (b) * The fire in the apartment were caused by a cigarette butt thrown on the carpet.
   (c) Those fires in the apartment were caused by a cigarette butt thrown on the carpet.
   (d) * That fire in the apartment were caused by a cigarette butt thrown on the carpet.

2. (a) The boxes for the toy were found in the backyard.
   (b) * The box for the toy were found in the backyard.
   (c) Those boxes for the toy were found in the backyard.
   (d) * That box for the toy were found in the backyard.

3. (a) The illustrations in the manual were done by a well-known artist.
   (b) * The illustration in the manual were done by a well-known artist.
   (c) Those illustrations in the manual were done by a well-known artist.
   (d) * That illustration in the manual were done by a well-known artist.

4. (a) The addresses on the envelope were not legible at all.
   (b) * The address on the envelope were not legible at all.
   (c) Those addresses on the envelope were not legible at all.
   (d) * That address on the envelope were not legible at all.

5. (a) The definitions in the dictionary were not helpful for understanding the word.
   (b) *The definition in the dictionary were not helpful for understanding the word.
   (c) Those definitions in the dictionary were not helpful for understanding the word.
   (d) *That definition in the dictionary were not helpful for understanding the word.

6. (a) The badges on the uniform were made in China.
   (b) *The badge on the uniform were made in China.
   (c) Those badges on the uniform were made in China.
   (d) *That badge on the uniform were made in China.

7. (a) The stories in the magazine were unknown to her for many years.
   (b) *The story in the magazine were unknown to her for many years.
   (c) Those stories in the magazine were unknown to her for many years.
   (d) *That story in the magazine were unknown to her for many years.

8. (a) The drawings in the textbook were much better in this edition.
   (b) *The drawing in the textbook were much better in this edition.
   (c) Those drawings in the textbook were much better in this edition.
   (d) *That drawing in the textbook were much better in this edition.

9. (a) The doors to the office were left unlocked by the cleaning service.
   (b) *The door to the office were left unlocked by the cleaning service.
   (c) Those doors to the office were left unlocked by the cleaning service.
   (d) *That door to the office were left unlocked by the cleaning service.

10. (a) The memos on the board were about the delinquent tax return.
    (b) *The memo on the board were about the delinquent tax return.
    (c) Those memos on the board were about the delinquent tax return.
    (d) *That memo on the board were about the delinquent tax return.

11 (a) The proposals for the project were under consideration for a long time.
   (b) *The proposal for the project were under consideration for a long time.
   (c) Those proposals for the project were under consideration for a long time.
   (d) *That proposal for the project were under consideration for a long time.

12. (a) The bags for the purchase were left on the counter by the customer.
    (b) *The bag for the purchase were left on the counter by the customer.
    (c) Those bags for the purchase were left on the counter by the customer.
    (d) *That bag for the purchase were left on the counter by the customer.

13. (a) The songs in the play were composed by a German musician.
    (b) *The song in the play were composed by a German musician.
    (c) Those songs in the play were composed by a German musician.
    (d) *That song in the play were composed by a German musician.

14. (a) The answers to the question were simpler than we had expected.
    (b) *The answer to the question were simpler than we had expected.
    (c) Those answers to the question were simpler than we had expected.
    (d) *That answer to the question were simpler than we had expected.

15. (a) The reasons for the test were to make sure the effect was reliable.
    (b) *The reason for the test were to make sure the effect was reliable.
    (c) Those reasons for the test were to make sure the effect was reliable.
    (d) *That reason for the test were to make sure the effect was reliable.

16. (a) The designs of the study were shown to be problematic in subsequent tests.
    (b) *The design of the study were shown to be problematic in subsequent tests.
    (c) Those designs of the study were shown to be problematic in subsequent tests.
    (d) *That design of the study were shown to be problematic in subsequent tests.

17. (a) The roads to the house were covered with water and mud.
    (b) *The road to the house were covered with water and mud.
    (c) Those roads to the house were covered with water and mud.
    (d) *That road to the house were covered with water and mud.

18. (a) The words on the screen were hard to recognize.
    (b) *The word on the screen were hard to recognize.
    (c) Those words on the screen were hard to recognize.
    (d) *That word on the screen were hard to recognize.

19. (a) The causes of the accident were under investigation by the local police.
    (b) *The cause of the accident were under investigation by the local police.
    (c) Those causes of the accident were under investigation by the local police.
    (d) *That cause of the accident were under investigation by the local police.

20. (a) The balloons for the party were bigger than we thought.
    (b) *The balloon for the party were bigger than we thought.
    (c) Those balloons for the party were bigger than we thought.
    (d) *That balloon for the party were bigger than we thought.

Filler items: *where*-clause

(a) *Where*-Plausible (control), (b) *Where*-Implausible (control)

Note: The *Where*-Implausible (control) condition sentences are not themselves implausible, but use the same verbs as the Gapping-Implausible condition sentences (see above) as a control.

1. (a) Henry ordered coffee and tea at the cafe where sandwiches and cake were very popular.
   (b) Henry drank coffee and tea at the cafe where sandwiches and cake were very popular.

2. (a) Andrew heated brownies and cake at home where milk and coffee were gone.
   (b) Andrew baked brownies and cake at home where milk and coffee were gone.

3. (a) Jenny prepared the apples and the onions in the kitchen where the spoons and the forks were kept.
   (b) Jenny peeled the apples and the onions in the kitchen where the spoons and the forks were kept.

4. (a) Sam replaced the doors and the windows in the bedroom where the fans and the chairs were located.
   (b) Sam opened the doors and the windows in the bedroom where the fans and the chairs were located.

5. (a) Jason grabbed the pill and the water at the hospital where the money and the card mysteriously disappeared.
   (b) Jason swallowed the pill and the water at the hospital where the money and the card mysteriously disappeared.

6. (a) Sara drew the spider and the ant in the park where the car and the truck were parked.
   (b) Sara killed the spider and the ant in the park where the car and the truck were parked.

Filler items: *Do So* Anaphora

(a) Do So Anaphora-Adjunct, (b) ?Do So Anaphora-Argument, (c) Control-Adjunct, (d) Control-Argument

1. (a) Robin read a book on the couch and Leslie did so on the bench.
   (b) ? Robin put a book on the couch and Leslie did so on the bench.
   (c) Robin read a book on the couch and Leslie read a book on the bench.
   (d) Robin put a book on the couch and Leslie put a book on the bench.

2. (a) Sam wrote a letter on the desk and Tina did so on the table.
   (b) ? Sam placed a letter on the desk and Tina did so on the table.
   (c) Sam wrote a letter on the desk and Tina wrote a letter on the table.
   (d) Sam placed a letter on the desk and Tina placed a letter on the table.

3. (a) Eric folded the towel in the bathroom and Jessy did so in the bedroom.
   (b) ? Eric threw the towel in the bathroom and Jessy did so in the bedroom.
   (c) Eric folded the towel in the bathroom and Jessy folded the towel in the bedroom.
   (d) Eric threw the towel in the bathroom and Jessy threw the towel in the bedroom.

4. (a) Ben unfolded boxes in the truck and Hannah did so in the car.
   (b) ? Ben loaded boxes in the truck and Hannah did so in the car.
   (c) Ben unfolded boxes in the truck and Hannah unfolded boxes in the car.
   (d) Ben loaded boxes in the truck and Hannah loaded boxes in the car.

Appendix M:

Mean Raw Reading Times (Standard Deviations) per Segment and Condition

Segment / Condition	1	2	3	4	5	6	7	8 critical region	9 spill-over region	10
	Bill	ordered /drank	coffee	and tea	at the cafe,	and	Jane	[e] sandwiches /did [e]	and cake /too	at the bakery./ with his brother.
L1-English										
(a) Gapping-P	437.73 (130.37)	429.68 (137.26)	449.55 (133.98)	519.38 (188.04)	545.16 (167.72)	455.11 (100.18)	404.04 (112.93)	407.68 (127.26)	501.37 (154.15)	526.00 (144.98)
(b) *Gapping-I	449.44 (134.77)	424.62 (126.52)	451.85 (135.11)	504.01 (157.73)	536.56 (153.59)	448.60 (94.96)	401.68 (110.22)	411.85 (129.28)	538.53 (173.54)	543.35 (150.59)
(c) VPE-P (baseline)	423.73 (124.88)	402.32 (124.60)	431.07 (141.46)	509.94 (171.97)	562.70 (161.32)	452.31 (106.71)	392.15 (104.80)	385.82 (94.67)	374.77 (92.62)	525.87 (153.99)
(d) VPE-I (baseline)	429.23 (119.97)	415.74 (128.98)	439.54 (128.89)	496.21 (151.83)	542.67 (161.02)	460.92 (102.46)	402.77 (101.90)	388.66 (92.85)	377.83 (98.53)	562.56 (160.18)
L2-English										
(a) Gapping-P	469.52 (159.72)	481.95 (167.70)	552.27 (219.41)	715.92 (286.49)	658.03 (233.29)	475.05 (141.73)	425.69 (148.31)	491.38 (181.18)	658.47 (243.44)	610.06 (247.14)
(b) *Gapping-I	453.42 (170.15)	476.21 (182.12)	543.57 (237.67)	642.30 (271.61)	667.99 (247.90)	474.24 (139.30)	437.25 (155.91)	467.57 (172.76)	743.64 (278.14)	661.13 (255.29)
(c) VPE-P (baseline)	455.30 (172.95)	477.65 (185.19)	554.70 (225.50)	688.59 (299.02)	663.06 (243.82)	444.73 (108.40)	421.88 (133.17)	398.92 (126.01)	367.62 (114.91)	648.14 (256.90)
(d) VPE-I (baseline)	438.28 (139.95)	456.13 (181.44)	563.46 (229.22)	643.14 (270.92)	672.35 (250.36)	460.16 (136.63)	435.01 (155.15)	394.76 (115.16)	364.31 (92.00)	600.23 (257.22)

Notes. L1-English: English native speakers; L2-English: L1-Korean L2ers of English.
The VPE-Implausible condition sentences are not themselves implausible, but use the same verbs as the *Gapping-Implausible condition sentences as a control.
Gapping-P: Gapping-Plausible; *Gapping-I: Gapping-Implausible; VPE-P: VPE-Plausible; VPE-I: VPE-Implausible.

REFERENCES

Abdulkarim, L., & Roeper, T. (1997). Economy of representation: Ellipsis and NP reconstructions. In A. Sorace, C. Heycock & R. Schillcock (Eds.), *Proceedings of GALA 97 Conference on Language Acquisition*. Edinburgh, UK: University of Edinburgh.

Abrahamsson, N. (2012). Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition*, *34*, 187–214.

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249–306.

Agbayani, B., & Zoerner, E. (2004). Gapping, pseudogapping and sideward movement. *Studia Linguistica*, *58*, 185–211.

Ahn, H.-D. (2018, June). *On relative clause deletion in English and Korean*. Talk presented at the 2018 Linguistic Society of Korea (LSK) Conference, Kyung Hee University, Seoul, Korea.

Al-Thubaiti, K. A. (2019). Selective vulnerability in very advanced L2 grammars: Evidence from VPE constraints. *Second Language Research*, *35*, 225–252.

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1–L2 similarity and structural distance. *Second Language Research*, *30*, 275–306.

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition, 106*, 87–129.

Arregui, A., Clifton, C., Frazier, L., & Moulton, K. (2006). Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language*, *55*, 232–246.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Berwick, R., & Weinberg, A. (1984). *The grammatical basis of linguistic performance: Language use and acquisition*. Cambridge, MA: MIT Press.

Biondo, N., Vespignani, F., & Dillon, B. (2019). Attachment and concord of temporal adverbs: Evidence from eye movements. *Frontiers in Psychology*, *10*, 983.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the Natural Language Toolkit*. Sebastopol, CA: O'Reilly.

Bley-Vroman, R. W. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41–68). Cambridge, UK: Cambridge University Press.

Bley-Vroman, R. W. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, *20*, 3–49.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer (Version 6.0.36) [Software]. Available from http://www.praat.org/

Bos, J., & Spenader, J. (2011). An annotated corpus for the analysis of VP ellipsis. *Language Resources and Evaluation*, *45*, 463–494.

Boxell, O., & Felser, C. (2017). Sensitivity to parasitic gaps inside subject islands in native and nonnative sentence processing. *Bilingualism: Language and Cognition*, *20*, 494–511.

Braunwald, S. R. (1997). The development of because and so: Connecting language, thought and social understanding. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships in the production and comprehension of text*. Hillsdale, NJ: Lawrence Erlbaum Associate.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, *64*, 311–317.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77–126.

Carlson, K. (2001). The effects of parallelism and prosody in the processing of gapping structures. *Language and Speech*, *44*, 1−26.

Carlson, K. (2002). *Parallelism and prosody in the processing of ellipsis sentences (Outstanding dissertations in linguistics series)*. New York: Routledge.

Carlson, K., Dickey, M. W., & Kennedy, C. (2005). Structural economy in the processing and representation of gapping sentences. *Syntax*, *8*, 208−228.

Carretero, L. L. (1999). VP-ellipsis in English and Spanish and the features of auxiliaries. *Probus*, *11*, 263–298.

Carroll, S. E. (2001). *Input and evidence: The raw material of second language acquisition*. Philadelphia, PA: John Benjamins.

Casasanto, L. S., Hofmeister, P., Sag, I. A., Ohlsson, S., & Catrambone, R. (2010). Understanding acceptability judgments: Additivity and working memory effects. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 224–229). Austin, TX: Cognitive Science Society.

Chao, W. (1987). *On ellipsis* (Unpublished doctoral dissertation). University of Massachusetts, Amherst, MA.

Cho, S. W. (2001). *Missing objects and the semantic representation in Korean discourse*. Talk presented at the Ninth Harvard International Symposium on Korean Linguistics, Harvard University, Cambridge, Massachusetts.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N., & Lasnik, H. (1993). Principles and parameters theory. In J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann (Eds.), *Syntax: An international handbook of contemporary research* (pp. 506–569). Berlin: Walter de Gruy.

Chung, D. (2004). A multiple dominance analysis of right node sharing constructions. *Language Research*, *40*, 791–811.

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42.

Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*, 564–570.

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*, 693–706.

Claus, B. (2015). Verb gapping: An action-gap compatibility study. *Acta psychologica*, *156*, 104–113.

Cole, P. (1987). Null objects in universal grammar. *Linguistic Inquiry*, *18*, 597–612.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type-token ratio (MATTR). *Journal of quantitative linguistics*, *17*, 94–100.

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, *14*, 597–650.

Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax*. Oxford: Oxford University Press.

Cunnings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*, 659–678.

Dagnac, A. (2010). Modal ellipsis in French, Spanish and Italian: Evidence for a TP-deletion analysis. In K. Arregi, Z. Fagyal, S. A. Montrul & A. Tremblay (Eds.), *Romance linguistics 2008: Interactions in Romance: Selected papers from the 38th Linguistic Symposium on Romance Languages* (pp. 157–170). Amsterdam: John Benjamins.

Dallas, A. C. (2008). *Influences of verbal properties on second-language filler-gap resolution: A cross-methodological study* (Unpublished doctoral dissertation). University of Florida, Gainesville, FL.

Dalrymple, M., Stuart, M. S., & Fernando, C. N. P. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy*, *14*, 399–452.

Davies, M. (2008–). The corpus of contemporary American English: 450 million words, 1990–present. Available from http://corpus.byu.edu/coca

de Vries, M. (2017). Across-the-board phenomena. In M. Everaert & H. van Riemsdijk (Eds.), *The Wiley Blackwell companion to syntax* (pp. 1–31). Oxford, UK: Blackwell.

Deen, K., Bondoc, I., Camp, A., Estioca, S., Hwang, H., Shin, G.-H., Takahashi, M., Zenker, F., & Zhong, J. (2018). Repetition brings success: Revealing knowledge of the passive voice. In A. B. Bertolini & M. J. Kaplan (Eds.), *Proceedings of the 42nd Annual Boston University Conference on Language Development* (pp. 200–213). Somerville, MA: Cascadilla Press.

Dekeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean?, In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 88–108). New York: Oxford University Press.

Dekydtspotter, L., & Renaud, C. (2014). On second language processing and grammatical development: The parser in second language acquisition. *Linguistic approaches to bilingualism*, *4*, 131–165.

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. G. O'Brien, C. Shea & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 33–40). Somerville, MA: Cascadilla Proceedings Project.

Diehl, J. J., Friedberg, C., Paul, R., & Snedeker, J. (2014). The use of prosody during syntactic processing in children and adolescents with autism spectrum disorders. *Development and Psychopathology*, *27*, 867–884.

Drummond, A. (2007). Ibex, experiment platform. http://spellout.net/ibexfarm/

Duffield, N. G., & Matsuo, A. (2009). Native speakers' versus L2 learners' sensitivity to parallelism. *Studies in Second Language Acquisition*, *31*, 93–123.

Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives; how cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language*, *36*, 829–854.

Felser, C. (2019). Structure-sensitive constraints in non-native sentence processing. *Journal of the European Second Language Association*, *3*, 12–22.

Felser, C., Cunnings, I., Batterham, C., & Clahsen, H. (2012). The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition*, *34*, 67–98.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, *23*, 9–36.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*, 453–489.

Fernandez, L., Höhle, B., Brock, J., & Nickels, L. (2018). Investigating auditory processing of syntactic gaps with L2 speakers using pupillometry. *Second Language Research*, *34*, 201–227.

Ferreira, F., & Clifton, C. Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Fiengo, R., & May, R. (1994). *Indices and identity*. Cambridge, MA: MIT Press.

Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguistic research*, *27*, 339–374.

Fodor, J. D., & Sakas, W. G. (2017). Learnability. In I. Roberts (Ed.), *The oxford handbook of Universal Grammar* (pp. 249–269). Oxford: Oxford University Press.

Foley, C., Núñez del Prado, Z., Barbier, I., & Lust, B. (2003). Knowledge of variable binding in VP-ellipsis: Language acquisition research and theory converge. *Syntax*, *6*, 52–83.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies* (Unpublished doctoral dissertation). University of Connecticut, Storrs, CT.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559 –586). Hillsdale, NJ: Erlbaum.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, *18*, 51–60.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry*, *25*, 407–454.

Goldberg, L. M. (2005). *Verb-stranding VP ellipsis: A cross-linguistic study* (Unpublished doctoral dissertation). McGill University, Quebec, Canada.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Guiraud, P. (1954). *Les charactères statistiques du vocabulaire: Essai de méthodologie* [The statistical characteristics of vocabulary: An essay in methodology]. Paris: Presses Universitaires de France.

Günther, C. (2012). *The elliptical noun phrase in English: Structure and use*. New York: Routledge.

Ham, S., Lee, Y., Lee, D., Choi, J., Kim, J., Kim, H., Myung, E., Park, S., Ahn, S., Eun, J., Lee, D., Im, N., Jun, M., & Jun, U. (2011). *Elementary school English textbook 3–4*. Seoul, Korea: Cheonjae Gyoyook.

Ham, S., Lee, Y., Lee, D., Choi, J., Kim, J., Kim, H., Myung, E., Park, S., Ahn, S., Eun, J., Lee, D., Im, N., Jun, M., & Jun, U. (2012). *Elementary school English textbook 5–6*. Seoul, Korea: Cheonjae Gyoyook.

Hankamer, J., & Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry, 7*, 391–428.

Hardt, D. (1993). *Verb phrase ellipsis: Form, meaning, and processing* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA.

Hardt, D. (1997). An empirical approach to VP ellipsis. *Computational Linguistics*, *23*, 525–541.

Harris, J. A., & Carlson, K. (2018). Information structure preferences in focus-sensitive ellipsis: How defaults persist. *Language and Speech*, *61*, 480–512.

Hawkins, R. (2012). Knowledge of English verb phrase ellipsis by speakers of Arabic and Chinese. *Linguistic Approaches to Bilingualism*, *2*, 404–438.

Haznedar, B. (2013). Child second language acquisition from a generative perspective. *Linguistic Approaches to Bilingualism*, *3*, 26–47.

Herbay, A. C., Gonnerman, L. M., & Baum, S. R. (2018). How do French–English bilinguals pull verb particle constructions off? Factors influencing second language processing of unfamiliar structures at the syntax-semantics interface. *Frontiers in psychology*, *9*, 1885.

Hestvik, A. (1995). Reflexives and ellipsis. *Natural Language Semantics*, *3*, 211–237.

Hoeks, J. C., Redeker, G., & Hendriks, P. (2009). Fill the gap! Combining pragmatic and prosodic information to make gapping easy. *Journal of psycholinguistic research*, *38*, 221–235.

Hofmann, M. (2006). *Verarbeitung elliptischer Satzkonstruktionen beim Sprachverstehen* [Processing of elliptical constructions in speech comprehension] (Unpublished doctoral dissertation). Universität Hamburg, Hamburg, Germany.

Hoji, H. (1998). Null object and sloppy identity in Japanese. *Linguistic Inquiry*, *29*, 127–152.

Honnibal, M., &, Montani, I. (to appear). spaCy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. Available from https://spacy.io/

Hopp, H. (2007*). Ultimate attainment at the interfaces in second language acquisition: Grammar and processing* (Unpublished doctoral dissertation). University of Groningen, Groningen, Netherlands.

Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition*, *12*, 463–483.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*, 901–931.

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, *17*, 5–27.

Houser, M. J. (2010). *The syntax and semantics of do so anaphora* (Unpublished doctoral dissertation). University of California, Berkeley, Berkeley, CA.

Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, *15*, 531–574.

Huang, C.-T. J. (1988). Comments on Hasegawa's paper. In T. Wako & M. Nakayama (Eds.), *Proceedings of Japanese Syntax Workshop: Issues on Empty Categories* (pp. 77–93). New London, CT: Connecticut College.

Huang, C.-T. J. (1991). Remarks on the status of the null object. In R. Freidin (Ed.), *Principles and parameters in comparative grammar* (pp. 56–76). Cambridge, MA: MIT Press.

Hunt, K. W. (1966). Recent measures in syntactic development. *Elementary English*, *43*, 732–739.

Hyams, N. M. (1986). *Language acquisition and the theory of parameters*. Dordrecht, Netherlands: Reidel.

Hyams, N. M. (2011). Missing subjects in early child language. In J. de Villiers & T. Roeper (Eds.), *Handbook of language acquisition theory in generative grammar* (pp. 13–52). Dordrecht, Netherlands: Kluwer.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539–587). Malden, MA: Blackwell.

Jackendoff, R. (1971). Gapping and related rules. *Linguistic Inquiry*, *2*, 21–36.

Jackson, C. N., & van Hell, J. G. (2011). The effects of L2 proficiency level on the processing of wh-questions among Dutch second language speakers of English. *International Review of Applied Linguistics in Language Teaching*, *49*, 195–219.

Jaeggli, O., & Hyams, N. M. (1988). Morphological uniformity and the setting of the null subject parameter. In J. Blevins & J. Carter (Eds.), *Proceedings of North-Eastern Linguistic Society 18* (pp. 238–253). Amherst, MA: University of Massachusetts.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.

Jiang, N., Hu, G., Chrabaszcz, A., & Ye, L. (2017). The activation of grammaticalized meaning in L2 processing: Toward an explanation of the morphological congruency effect. *International Journal of Bilingualism*, *21*, 81–98.

Johannessen, J. (1996). Partial agreement and coordination. *Linguistic Inquiry*, *27*, 661–675.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Johnson, K. (2000). Gapping determiners. In K. Schwabe & N. Zhang (Eds.), *Ellipsis in conjunction* (pp. 95–115). Tübingen, Germany: Niemeyer.

Johnson, K. (2006). Gapping. In M. Everaert & H. C. van Riemsdijk (Eds.), *The Blackwell companion to syntax* (Vol. 2, pp. 407–435). Cambridge, MA: Blackwell.

Johnson, K. (2009). Gapping is not (VP-) ellipsis. *Linguistic Inquiry*, *40*, 289–328.

Johnson, K. (2019). Gapping and stripping. In J. van Craenenbroeck & T. Temmerman (Eds.), *The Oxford handbook of ellipsis* (pp. 562–604). Oxford: Oxford University Press.

Johnson, V. E., de Villiers, J. G., & Seymour, H. N. (2005). Agreement without understanding? The case of third person singular /s/. *First Language*, *25*, 317–330.

Juffs, A., & Rodriguez, G. A. (2014). *Second Language Sentence Processing*. New York, NY: Routledge.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238.

Kaan, E., Ballantyne, J., & Wijnen, F. (2015). Effects of reading speed on second-language sentence processing. *Applied Psycholinguistics*, *36*, 799–830.

Kaan, E., Overfelt, C., Tromp, D., & Wijnen, F. (2013). Processing gapped verbs. *Journal of Psycholinguistic Research*, *42*, 307–338.

Kaan, E., Wijnen, F., & Swaab, T. Y. (2004). Gapping: Electrophysiological evidence for immediate processing of "missing" verbs in sentence comprehension. *Brain and Language*, *89*, 584–592.

Kanno, K. (1999). Acquisition of verb gapping in Japanese by Mandarin and English speakers. In K. Kanno (Ed.), *The acquisition of Japanese as a second language* (pp. 159–173). Amsterdam: John Benjamins.

Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition*, *37*, 1–32.

Keating, G. D., Jegerski, J., & VanPatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, *19*, 36–49.

Kehler, A. (2000). Coherence and the resolution of ellipsis. *Linguistics and Philosophy*, *23*, 533–575.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar.* Stanford, CA: CSLI.

Kidd, E., Stewart, A., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, *38*, 222–234.

Kim, J. (2012). *Comprehension of elided phrases in Korean and English: VP-ellipsis, null object constructions, and one-substitution* (Unpublished doctoral dissertation). University of Hawai‘i at Mānoa, Honolulu, HI.

Kim, J. (2015). Korean L2 Learners' comprehension of VP-ellipsis and one-substitution in English. *Enekwahakyenkwu* [The Journal of Linguistics Science], *75*, 107–128.

Kim, J.-S. (1997). What syntactic focus movement tells about VP Ellipsis in Korean and Japanese. *Ene* [Language], *22*, 433–452.

Kim, J.-S. (2006). On the size of ellipsis in Korean focus constructions. *Inmwunyenkwu* [Journal of the Humanities], *50*, 271–302.

Kim, J. W., Ahn, B., Oh, J., Kim, B., Ko, E., Ahn, K., Lee, J., Choi, Y., Han, S., Kim, J., & Yin, J. (2013). *High school English 1–2*. Seoul, Korea: Cheonjae Gyoyook.

Kim, J. W., Ahn, B., Oh, J., Kim, S., Park, S., Shin, S., Jo, S., Jo, H., Kim, B., & Yin, J. (2013). *Middle school English 1–3*. Seoul, Korea: Cheonjae Gyoyook.

Kim, J. W., Ahn, B., Oh, J., Shin, M., Hong, J., & Jang, Y. (2013). *High school basic English.* Seoul, Korea: Cheonjae Gyoyook.

Kim, K. (2014). *Unveiling linguistic competence by facilitating performance* (Unpublished doctoral dissertation). University of Hawai‘i at Mānoa, Honolulu, HI.

Kim, M. (2010). How do native English-speaking teachers perceive co-teaching? *Ungyongenehak* [Applied Linguistics], *26*, 213–249.

Kim, N., Carlson, K., Dickey, M., & Yoshida, M. (2020). Processing gapping: Parallelism and grammatical constraints. *Quarterly Journal of Experimental Psychology*, *73*, 781–798.

Kim, S. (1999). Sloppy/strict identity, empty objects, and NP ellipsis. *Journal of East Asian Linguistics*, *8*, 255–284.

Kitaev, N., & Klein, D. (2018). Constituency parsing with a self-attentive encoder. In I. Gurevych & Y. Miyao (Eds.), *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 2676–2686). Melbourne, Australia: Association for Computational Linguistics.

Kliegl, R., & Laubrock, J. (2018). Eye-movement tracking during reading. In A. M. de Groot & P. Hagoort (Eds.), *Research methods in psycholinguistics and the neurobiology of language: A practical guide* (pp. 68–88). Oxford: John Wiley & Sons.

Kobayashi, A. (2005). Constituency in gapping constructions. *Linguistic Analysis*, *32*, 184–225.

Koornneef, A. W., & Van Berkum, J. J. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, *54*, 445–465.

Kuiken, F., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 143–170). Amsterdam: John Benjamins.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lasnik, H. (1995). Verbal morphology: Syntactic structures meets the Minimalist Program. In H. Campos & P. Kempchinksy (Eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero* (pp. 251–275). Washington, D.C.: Georgetown University Press.

Lasnik, H. (1999). *Minimalist analysis*. Malden, MA: Blackwell.

Lee, J., Kim, H., Ha, M., Kim, S., Kim, E., Kwon, M., Yoon, K., Kim, D., Yoon, S., & Park, Y. (2011). *Elementary school English textbook 3–4*. Seoul, Korea: Daegyo.

Lee, J., Kim, H., Ha, M., Kim, S., Kim, E., Kwon, M., Yoon, K., Kim, D., Yoon, S., & Park, Y. (2012). *Elementary school English textbook 5–6*. Seoul, Korea: Daegyo.

Legendre, G., Culbertson, J., Zaroukian, E., Hsin, L., Barrière, I., & Nazzi, T. (2014). Is children's comprehension of subject–verb agreement universally late? Comparative evidence from French, English, and Spanish. *Lingua*, *144*, 21–39.

Lenneberg, E. (1967). *Language in the context of growth and maturation*. New York: John Wiley & Sons.

Li, H.-J. G. (1998). Null object and VP ellipsis in Chinese. In H. Lin (Ed.), *Proceedings of the 9th North American Conference on Chinese Linguistics* (pp. 151–172). Los Angeles, CA: GSIL, University of Southern California.

Li, P. (2000). The acquisition of lexical and grammatical aspect in a self-organizing feature-map model. In L. Gleitman & A. Jashi (Eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 304–309). Mahwah, NJ: Lawrence Erlbaum.

Lobeck, A. (1995). *Ellipsis: Functional heads, licensing, and identification*. New York: Oxford University Press.

Lobeck, A. (2008). Ellipsis in DP. In M. Everaert & H. C. Van Riemsdijk (Eds.), *The Blackwell companion to syntax*. Malden, MA: Blackwell Publishing.

Long, M. H. (2005). Problems with supposed counter-evidence to the critical period hypothesis. *International Review of Applied Linguistics*, *43*, 287–317.

Lopez, L. (1994). The syntactic licensing of VP-ellipsis: A comparative study of Spanish and English. In M. L. Mazzola (Ed.), *Issues and theory in Romance linguistics: Selected papers from the Linguistic Symposium on Romance Languages XXIII* (pp. 333–354). Washington, DC: Georgetown University Press.

MacWhinney, B. (1991). *The CHILDES project: Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78.

Matsuo, A. (2007). Differing interpretations of empty categories in English and Japanese VP ellipsis contexts. *Language Acquisition*, *14*, 3–29.

Matsuo, A., & Duffield, N. (2001). VP-ellipsis and anaphora in child language acquisition. *Language Acquisition*, *9*, 301–327.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*, 381–401.

Merchant, J. (2001). *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford, UK: Oxford University Press.

Merchant, J. (2003). *Remarks on stripping* (Unpublished manuscript). Chicago, IL: University of Chicago.

Merchant, J. (2004). Fragments and ellipsis. *Linguistics and Philosophy*, *27*, 661–738.

Ministry of Education. (2015). *The English national curriculum*. Seoul, Korea: Ministry of Education.

Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First language*, *39*, 527–546.

Mukai, E. (2003). On verbless conjunction in Japanese. In M. Kadowaki & S. Kawahara (Eds.), *Proceedings of the North East Linguistic Society 33* (pp. 205–223). Amherst, MA: GLSA, University of Massachusetts, Amherst.

Murphy, G. L. (1985a). Processes of understanding anaphora. *Journal of Memory and Language*, *24*, 290–303.

Murphy, G. L. (1985b). Psychological explanations of deep and surface anaphora. *Journal of Pragmatics*, *9*, 171–198.

Newman, A., Tremblay, A., Nichols, E., Neville, H., & Ullman, M. (2012). The influence of language proficiency on lexical-semantic processing in native and late learners of English: ERP evidence. *Journal of Cognitive Neuroscience*, *24*, 1205–1223.

Nielsen, L. A. (2005). *A corpus-based study of verb phrase ellipsis identification and resolution* (Unpublished doctoral dissertation). King's College London, London, UK.

O'Grady, W. (1999). Gapping and coordination in second language acquisition. In K. Kanno (Ed.), *The acquisition of Japanese as a second language* (pp. 141–157). Amsterdam: John Benjamins.

Omaki, A., Lau, E. F., White, I. D., Dakan, M. L., Apple, A., & Phillips, C. (2015). Hyper-active gap filling. *Frontiers in psychology*, *6*, 384.

Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition*, *33*, 563–588.

Orfitelli, R., & Hyams, N. M. (2012). Children's grammar of null subjects: Evidence from comprehension. *Linguistic inquiry*, *43*, 563–590.

Otani, K., & Whitman, J. (1991). V-raising and VP-ellipsis. *Linguistic Inquiry*, *22*, 345–358.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition*, *25*, 501–528.

Park, K.-S. (2014). *Information structure and dative word-order alternations in English and Korean: L1 children, L2 children, and L2 adults* (Unpublished doctoral dissertation). University of Hawai‘i at Mānoa, Honolulu, HI.

Park, M.-K. (1997). The syntax of VP ellipsis in Korean. *Ehakyenkwu* [Language Research], *33*, 629–648.

Park, M.-K. (2015). Extraction out of overt anaphora: Korean *kule(h)* ‘so’ versus English *so*. *Eneyenkwu* [Linguistic Research], *32*, 693–718.

Peirce, J. (2017). PsychoPy (Version 1.85.3) [Software]. Available from http://www.psychopy.org/

Peristeri, E., & Tsimpli, I. M. (2013). Pronoun processing in Broca’s aphasia: Discourse–syntax effects in ambiguous anaphora resolution. *Aphasiology*, *27*, 1381–1407.

Python Software Foundation. (2018). Python language reference (Version 3.7.1) [Software]. Available from http://www.python.org

R Core Team. (2018). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Available from https://www.R-project.org/

Repp, S. (2009). Negation in Gapping. Oxford: Oxford University Press.

Roberts, L., & Liszka, S. A. (2020). Grammatical aspect and L2 learners’ online processing of temporarily ambiguous sentences in English: A self-paced reading study with German, Dutch and French L2 learners. *Second Language Research*. Advanced online publication. https://doi.org/10.1177/0267658319895551

Ross, J. R. (1970). Gapping and the order of constituents. In M. Bierwisch & K. E. Heidolph (Eds.), *Progress in linguistics* (pp. 249–259). The Hague, Netherlands: Mouton.

Rouveret, A. (2010). *VP ellipsis, phases and the syntax of morphology* (Unpublished manuscript). Paris Diderot University, Paris, France.

Rouveret, A. (2012). VP ellipsis, phrases and the syntax of morphology. *Natural Language and Linguistic Theory*, *30*, 897–963.

Ruijgrok, B. J. (2018). *Tapping into semantic recovery: An event-related potential study on the processing of gapping and stripping* (Unpublished doctoral dissertation). Netherlands Graduate School of Linguistics (LOT), Amsterdam, Netherlands.

Sag, I. (1976). *Deletion and logical form* (Unpublished doctoral dissertation). Massachusetts Institution of Technology, Cambridge, MA.

Sag, I., & Hankamer, J. (1984). Towards a theory of anaphoric processing. *Linguistics and Philosophy*, *7*, 325–345.

Sakas, W. G., & Fodor, J. D. (2012). Disambiguating syntactic triggers. *Language Acquisition*, *19*, 83–143.

Santos, A. L. (2006). *Minimal answers: Ellipsis, syntax and discourse in the acquisition of European Portuguese* (Unpublished doctoral dissertation). Universidade de Lisboa, Lisbon, Portugal.

Schwartz, B. D. (1987). *The modular basis of second language acquisition* (Unpublished doctoral dissertation). University of Southern California, Los Angeles, CA.

Schwartz, B. D. (1992). Testing between UG-based and problem-solving models of L2A: Developmental sequence data. *Language Acquisition*, *2*, 1–19.

Schwartz, B. D. (1997). On the basis of the Basic Variety... *Second Language Research*, *13*, 386–402.

Schwartz, B. D. (1999). Let's make up your mind: "Special nativist" perspectives on language, modularity of mind, and nonnative language acquisition. *Studies in Second Language Acquisition*, *21*, 635–655.

Schwartz, B. D. (2004). Why child L2 acquisition? In J. van Kampen & S. Baauw (Eds.), *Proceedings of GALA 2003* (Vol. 1, pp. 47–66). Utrecht, Netherlands: Netherlands Graduate School of Linguistics (LOT).

Schwartz, B. D., & Sprouse, R. A. (1994). Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. D. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp. 317–368). Amsterdam: John Benjamins.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, *12*, 40–72.

Schwartz, B. D., & Sprouse, R. A. (2000). When syntactic theories evolve: Consequences for L2 acquisition research. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 156–186). Malden, MA: Blackwell.

Schwartz, B. D., & Sprouse, R. A. (2013). Generative approaches and the poverty of the stimulus. In J. Herschensohn & M. Young-Scholten (Eds.), *Handbook of second language acquisition* (pp. 137–158). Cambridge, UK: Cambridge University Press.

Sharwood Smith, M., & Truscott, J. (2014). *The multilingual mind: A modular processing perspective*. Cambridge: Cambridge University Press.

Slabakova, R. (2016). *Second language acquisition*. Oxford: Oxford University Press.

Smith, P. A. (2011). Attention, working memory, and grammaticality judgment in typical young adults. *Journal of Speech, Language, and Hearing Research*, 54, 918–931.

Sohn, H. M. (1999). *The Korean language*. Cambridge, UK: Cambridge University Press.

Song, H. S., & Schwartz, B. D. (2009). Testing the Fundamental Difference Hypothesis: L2 adult, L2 child, and L1 child comparisons in the acquisition of Korean wh-constructions with negative polarity items. *Studies in Second Language Acquisition*, *31*, 323–361.

Spooren, W., & Sanders, T. (2008). The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, *40*, 2003–2026.

Sprouse, R. A. (2011). The interface hypothesis and full transfer/full access/full parse: A brief comparison. *Linguistic approaches to bilingualism*, *1*, 97–100.

Streb, J., Hennighausen, E., & Rösler, F. (2004). Different anaphoric expressions are investigated by event-related brain potentials. *Journal of Psycholinguistic Research*, *33*, 175–201.

Stroik, T. (2001). On the light verb hypothesis. *Linguistic Inquiry*, *32*, 362–369.

Stroup, W. W. (2012). *Generalized linear mixed models: Modern concepts, methods and applications*. Boca Raton, FL: CRC Press.

Tanenhaus, M., & Carlson, G. N. (1990). Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes*, *5*, 257–280.

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in on-line L2 sentence comprehension. *Bilingualism: Language and Cognition*, *17*, 277–293.

Thornton, R., & Wexler, K. (1999). *Principle B, VP ellipsis, and interpretation in child grammar*. Cambridge, MA: MIT Press.

Trueswell, J, Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89–134.

Truscott, J., & Sharwood Smith, M. (2004) Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition*, 7, 1–20.

Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, *23*, 215–342.

Unsworth, S. (2005). *Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch* (Unpublished doctoral dissertation). Utrecht University, Utrecht, Netherlands.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of memory and language*, *55*, 157–166.

van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning, 62*, 148–171.

Villata, S., Tabor, W., & Franck, J. (2018). Encoding and retrieval interference in sentence comprehension: Evidence from agreement. *Frontiers in Psychology*, *9*, 2.

Wasow, T. (1972). *Anaphoric relations in English* (Unpublished doctoral dissertation). Massachusetts Institution of Technology, Cambridge, Massachusetts.

Wen, Z., Miyao, M., Takeda, A., Chu, W., & Schwartz, B. D. (2010). Proficiency effects and distance effects in non-native processing of English number agreement. In K. Franchini, K. Iserman & L. Keil (Eds.), *Proceedings of the 34th Annual Boston University Conference on Language Development* (pp. 445–456). Somerville, MA: Cascadilla.

Westergaard, M. (2014). Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation*, *14*, 26–45.

Westergaard, M. (2019). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research*. Advanced online publication. https://doi.org/10.1177/0267658319884116

White, L. (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied Linguistics*, *8*, 95–110.

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge, UK: Cambridge University Press.

Whong-Barr, M., & Schwartz, B. D. (2002). Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. *Studies in Second Language Acquisition*, *24*, 579–616.

Winkler, S. (2005). *Ellipsis and focus in generative grammar*. Berlin: de Gruyter.

Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics, 33*, 419–456.

Wulfeck, B. B. (1993). A reaction time study of grammaticality judgments in children. *Journal of Speech, Language, and Hearing Research*, *36*, 1208–1215.

Xu, L. (2003). Remarks on VP-ellipsis in disguise. *Linguistic Inquiry*, *34*, 163–171.

Zenker, F., & Kyle, K. (2019). *Investigating minimum text lengths for lexical diversity indices*. Manuscript submitted for publication.

Zheng, X., & Lemhöfer, K. (2019). The "Semantic P600" in second language processing: When syntax conflicts with semantics. *Neuropsychologia*, *127*, 131–147.

Zoerner, E. (1999). Left-peripheral deletion and gapping: The same but different. In M. M. T. Henderson (Ed.), *Papers of the Mid-America Linguistics Conference* (pp. 577–585). Lawrence, KS: University of Kansas.

Zukowski, A., & Larsen, J. (2011). *Wanna* contraction in children: Retesting and revising the developmental facts. *Language Acquisition*, *18*, 211–241.

Zwart, C. J.-W. (1997). *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch* (Vol. 39). Dordrecht, Netherlands: Springer Science & Business Media.